A.

```
                                        ATGTCCTCGGCGCCGTCTCCGGGG        24
                                         M  S  S  A  P  S  P  G

 25 ACTGGTTCGCCTCCATCTCCACCATCAAACTCCACAACCACCACTCCTCCTCCAGCTTCC    84
      T  G  S  P  P  S  P  P  S  N* S  T  T  T  T  P  P  P  A  S

 85 GCTCCTCCTCCCACCACACCTTCTTCTCCTCCGCCGCCATCCACTATTCCGACATCTCCT   144
      A  P  P  P  T  T  P  S  S  P  P  P  S  T  I  P  T  S  P

145 CCTCCTTCTTCTCGCTCTACACCTTCTGCTCCTCCTCCATCTCCACCAACTCCATCTACG   204
      P  P  S  S  R  S  T  P  S  A  P  P  P  S  P  P  T  P  S  T

205 CCGGGATCTCCACCTCCTCTTCCTCAGCCGTCTCCACCCGCTCCAACTACGCCCGGATCT   264
      P  G  S  P  P  P  L  P  Q  P  S  P  P  A  P  T  T  P  G  S

265 CCACCCGCACCTGTTACTCCTCCTACTCGAAACCCTCCACCTTCAGTCCCAGGACCACCG   324
      P  P  A  P  V  T  P  P· T  R  N  P  P  P  S  V  P  G  P  P

325 TCCAATCCTTCACGCGAAGGAGGATCTCCTCGACCTCCATCTTCTCCCTCGCCGCCGTCT   384
      S  N  P  S  R  E  G  G  S  P  R  P  P  S  S  P  S  P  P  S

385 CCTTCTTCCGACGGTTTATCAACAGGAGTGGTGGTGGGAATCGCCATCGGAGGAGTCGCT   444
      P  S  S  D  G  L  S  T  G  V  V  V  G  I  A  I  G  G  V  A

445 CTGCTTGTGATAGTGACTCTGATTTGTCTCCTCTGTAAGAAGAAACGACGGAGAGACGAA   504
      L  L  V  I  V  T  L  I  C  L  L  C  K  K  K  R  R  R  D  E

505 GAAGATGCTTACTATGTTCCTCCGCCACCTCCTCCTGGTCCCAAAGCCGGAGGACCTTAC   564
      E  D  A  Y  Y  V  P  P  P  P  P  G  P  K  A  G  G  P  Y

565 GGTGGACAGCAGCAACAATGGCGGCAACAAAACGCAACACCACCGTCAGATCATGTCGTG   624
      G  G  Q  Q  Q  Q  W  R  Q  Q  N  A  T  P  P  S  D  H  V  V

625 ACGTCACTACCACCACCACCTAAGGCTCCATCTCCACCACGGCAACCTCCTCCACCTCCA   684
      T  S  L  P  P  P  P  K  A  P  S  P  P  R  Q  P  P  P  P  P

685 CCACCGCCTTTCATGAGCAGCAGCGGCGGCTCCGACTACTCGGACCGTCCAGTTCTTCCT   744
      P  P  P  F  M  S  S  S  G  G  S  D  Y  S  D  R  P  V  L  P

745 CCACCGTCTCCAGGGCTTGTGTTAGGCTTCTCCAAAAGCACTTTCACATACGAGGAGCTA   804
      P  P  S  P  G  L  V  L  G  F  S  K  S  T  F  T  Y  E  E  L

805 GCTAGAGCCACCAATGGTTTCTCCGAGGCGAACTTGTTAGGACAAGGCGGGTTCGGTTAC   864
      A  R  A  T  N  I G  F  S  E  A  N  L  L  G  Q  G  G  F  G  Y

865 GTGCACAAAGGTGTGTTGCCTAGTGGGAAAGAAGTTGCTGTGAAGCAGTTGAAAGTTGGG   924
      V  H  K  G  V  L  P  S  II G  K  E  V  A  V  K  Q  L  K  V  G

925 AGTGGTCAGGGAGAGAGGGAGTTTCAGGCAGAGGTTGAGATCATCAGCAGAGTTCACCAC   984
      S  G  III Q  G  E  R  E  F  Q  A  E  V  E  L  L  I  S  R  V  IV H  H

985 AGGCATCTGGTGTCTCTTGTTGGTTATTGCATCGCCGGTGCCAAAAGATTGCTTGTCTAT  1044
      R  H  L  V  S  L  V  G  Y  C  I  A  V Q  A  K  R  L  L  L  V  Y

1045 GAGTTTGTTCCTAACAACAATCTCGAGCTTCACCTCCATGGCGAGGGACGGCCTACAATG  1104
      E  F  V  P  N  N  N  L  E  L  H  L  H  G  E  G  R  VI P  T  M

1105 GAATGGAGCACCAGATTGAAGATTGCTCTTGGATCTGCTAAAGGACTTTCTTATCTTCAT  1164
      E  W  S  T  R  L  K  I  A  L  G  S  A  K  G  L  S  Y  L  H

1165 GAAGATTGCAATCCTAAAATCATTCACCGTGATATCAAGGCTTCAAACATATTGATAGAT  1224
      E  D  C  N  P  K  I  I  H  R  D  I  K  A  S  N  I  L  I  D

1225 TTCAAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTTGCTAAGATTGCTTCTGATACAAAC  1284
      F  K  E  VII E  A  K  V  A  D  F  G  L  A  K  I  A  S  D  T  N

1285 ACGCATGTATCAACACGTGTGATGGGAACCTTTGGGTACTTGGCTCCGGAATACGCTGCA  1344
      T  H  V  VIII S  T  R  V  M  G  T  F  G  Y  L  A  P  E  Y  IX A  A

1345 AGCGGAAAGCTCACGGAGAAGTCTGACGTTTTCTCATTTGGCGTTGTGCTTTTGGAGCTC  1404
      S  G  K  L  T  E  K  S  D  V  F  S  F  G  V  V  L  L  E  L

1405 ATTACTGGACGTCGACCCGTTGATGCCAACAATGTCTATGTAGATGACAGCTTAGTTGAC  1464
      I  T  G  R  R  P  V  D  A  N  N  X V  Y  V  D  D  S  L  V  D

1465 TGGGCACGACCATTGCTTAACCGAGCATCTGAGCAAGGAGACTTTGAGGGTTTAGCTGAT  1524
      W  A  R  P  L  L  N  R  A  S  E  Q  G  D  F  E  G  L  A  D

1525 GCAAAGATGAATAATGGGTATGACAGAGAGGAGATGGCTCGCATGGTTGCTTGTGCTGCG  1584
      A  K  M  N  N  G  Y  D  R  E  XI E  M  A  R  M  V  A  C  A  A

1585 GCTTGTGTGTTCGCCATTCAGCTCGCCGCAGACCTCGCATGAGCCAGATTGTGCGTGCGTTA  1644
      A  C  V  R  H  S  A  R  R  P  R  M  S  Q  I  V  R  A  L

1645 GAAGGAAATGTATCACTGTCAGATCTTAACGAAGGGATGAGACCAGGTCAAAGCAATGTA  1704
      E  G  N* V  S  L  S  D  L  N  E  G  M  R  P  G  Q  S  N  V

1705 TACAGCTCATACGGAGGAAGCACCGATTATGACTCGAGCCAGTACAATGAAGACATGAAG  1764
      Y  S  S  Y  G  G  S  T  D  Y  D  S  S  Q  Y  N  E  D  M  K

1765 AAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTACAACGCCACGGGTGAGTACAGTAAT  1824
      K  F  R  K  M  A  L  G  T  Q  E  Y  N* A  T  G  E  Y  S  N

1825 CCGACCAGTGACTATGGACTGTACCCGTCTGGTTCAAGCAGCGAGGGCCAAACCACACGC  1884
      P  T  S  D  Y  G  L  Y  P  S  G  S  S  S  E  G  Q  T  T  R

1885 GAAATGGAGATGGGGAAGATTAAGAGAACCGGTCAGGGTTATAGTGGACCTTCTCTTTAA  1944
      E  M  E  M  G  K  I  K  R  T  G  Q  G  Y  S  G  P  S  L  •
```

Figure 1A

B.

Figure1B

```
 -95 ttaactctctggtctccgtgtctcctctcttctcctgctgcttccttttaacactctctt  -36
 -35 catttgccttttttgatttagatccaaagaagcagacATGTCCTCGGCGCCGTCTCCGGGG   24
                                          M  S  S  A  P  S  P  G
  25 ACTGGTTCGCCTCCATCTCCACCATCAAACTCCACAACCACCACTCCTCCTCCAGCTTCC   84
      T  G  S  P  P  S  P  P  S  N*  S  T  T  T  T  P  P  P  A  S
  85 GCTCCTCCTCCCACCACACCTTCTTCTCCTCCGCCGCCATCCACTATTCCGACATCTCCT  144
      A  P  P  P  T  T  P  S  S  P  P  P  P  S  T  I  P  T  S  P
 145 CCTCCTTCTTCTCGCTCTACACCTTCTGCTCCTCCTCCATCTCCACCAACTCCATCTACG  204
      P  P  S  S  R  S  T  P  S  A  P  P  P  S  P  P  T  P  S  T
 205 CCGGGATCTCCACCTCCTCTTCCTCAGCCGTCTCCACCCGCTCCAACTACGCCCGGATCT  264
      P  G  S  P  P  P  L  P  Q  P  S  P  P  A  P  T  T  P  G  S
 265 CCACCCGCACCTGTTACTCCTCCTACTCGAAACCCTCCACCTTCAGTCCCAGGACCACCG  324
      P  P  A  P  V  T  P  P  T  R  N  P  P  P  S  V  P  G  P  P
 325 TCCAATCCTTCACGCGAAGGAGGATCTCCTCGACCTCCATCTTCTCCCTCGCCGCCGTCT  384
      S  N  P  S  R  E  G  G  S  P  R  P  P  S  S  P  S  P  P  S
 385 CCTTCTTCCGACGGTTTATCAACAGGAGTGGTGGTGGGAATCGCCATCGGAGGAGTCGCT  444
      P  S  S  D  G  L  S  T  G  V  V  V  G  I  A  I  G  G  V  A
 445 CTGCTTGTGATAGTGACTCTGATTTGTCTCCTCTGTAAGAAGAAACGACGGAGAGACGAA  504
      L  L  I  V  T  L  I  C  L  L  C  K  K  K  R  R  R  D  E
 505 GAAGATGCTTACTATGTTCCTCCGCCACCTCCTCCTGGTCCCAAAGCCGGAGGACCTTAC  564
      E  D  A  Y  Y  V  P  P  P  P  P  P  G  P  K  A  G  G  P  Y
 565 GGTGGACAGCAGCAACAATGGCGGCAACAAAACGCAACACCACCGTCAGATCATGTCGTG  624
      G  G  Q  Q  Q  Q  W  R  Q  Q  N  A  T  P  P  S  D  H  V  V
 625 ACGTCACTACCACCACCACCTAAGGCTCCATCTCCACCACGGCAACCTCCTCCACCTCCA  684
      T  S  L  P  P  P  K  A  P  S  P  P  R  Q  P  P  P  P
 685 CCACCGCCTTTCATGAGCAGCAGCGGCGGCTCCGACTACTCGGACCGTCCAGTTCTTCCT  744
      P  P  P  F  M  S  S  G  G  S  D  Y  S  D  R  P  V  L  P
 745 CCACCGTCTCCAGGGCTTGTGTTAGGCTTCTCCAAAAGCACTTTCACATACGAGGAGCTA  804
      P  P  S  P  G  L  V  L  G  F  S  K  S  T  F  T  Y  E  E  L
 805 GCTAGAGCCACCAATGGTTTCTCCGAGGCGAACTTGTTAGGACAAGGCGGGTTCGGTTAC  864
      A  R  A  T  N  I  Q  F  S  E  A  N  L  L  G  Q  G  G  F  G  Y
 865 GTGCACAAAGGTGTGTTGCCTAGTGGGAAAGAAGTTGCTGTGAAGCAGTTGAAAGTTGGG  924
      V  H  K  G  V  L  P  S  II  G  K  E  V  A  V  K  Q  L  K  V  G
 925 AGTGGTCAGGGAGAGAGGGAGTTTCAGGCAGAGGTTGAGATCATCAGCAGAGTTCACCAC  984
      S  Q  III  Q  G  E  R  E  F  Q  A  E  V  E  I  I  S  R  V  IV  H  H
 985 AGGCATCTGGTGTCTCTTGTTGGTTATTGCATCGCCGGTGCCAAAAGATTGCTTGTCTAT  1044
      R  H  L  V  S  L  V  G  Y  C  I  A  V  Q  A  K  R  L  L  V  Y
1045 GAGTTTGTTCCTAACAACAATCTCGAGCTTCACCTCCATGGCGAGGGACGGCCTACAATG  1104
      E  F  V  P  N  N  N  L  E  L  H  L  H  G  E  G  R  VI  P  T  M
1105 GAATGGAGCACCAGATTGAAGATTGCTCTTGGATCTGCTAAAGGACTTTCTTATCTTCAT  1164
      E  W  S  T  R  L  K  I  A  L  G  S  A  K  G  L  S  Y  L  H
1165 GAAGATTGCAATCCTAAAATCATTCACCGTGATATCAAGGCTTCAAACATATTGATAGAT  1224
      E  D  C  N  P  K  I  I  H  R  D  I  K  A  S  N  I  L  I  D
1225 TTCAAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTTGCTAAGATTGCTTCTGATACAAAC  1284
      F  K  F  VII  E  A  K  V  A  D  F  G  L  A  K  I  A  S  D  T  N
1285 ACGCATGTATCAACACGTGTGATGGGAACCTTTGGGTACTTGGCTCCGGAATACGCTGCA  1344
      T  H  V  VIII  S  T  R  V  M  G  T  F  G  Y  L  A  P  E  Y  IX  A  A
1345 AGCGGAAAGCTCACGGAGAAGTCTGACGTTTTCTCATTTGGCGTTGTGCTTTTGGAGCTC  1404
      S  G  K  L  T  E  K  S  D  V  F  S  F  G  V  V  L  L  E  L
1405 ATTACTGGACGTCGACCCGTTGATGCCAACAATGTCTATGTAGATGACAGCTTAGTTGAC  1464
      I  T  G  R  R  P  V  D  A  N  N  X  V  Y  V  D  D  S  L  V  D
1465 TGGGCACGACCATTGCTTAACCGAGCATCTGAGCAAGGAGACTTTGAGGGTTTAGCTGAT  1524
      W  A  R  P  L  L  N  R  A  S  E  Q  G  D  F  E  G  L  A  D
1525 GCAAAGATGAATAATGGGTATGACAGAGAGGAGATGGCTCGCATGGTTGCTTGTGCTGCG  1584
      A  K  M  N  N  G  Y  D  R  E  XI  M  A  R  M  V  A  C  A  A
1585 GCTTGTGTTCGCCATTCAGCTCGCCGCAGACCTCGCATGAGCCAGATTGTGCGTGCGTTA  1644
      A  C  V  R  H  S  A  R  R  R  P  R  M  S  Q  I  V  R  A  L
1645 GAAGGAAATGTATCACTGTCAGATCTTAACGAAGGGATGAGACCAGGTCAAAGCAATGTA  1704
      E  G  N*  V  S  L  S  D  L  N  E  G  M  R  P  G  Q  S  N  V
1705 TACAGCTCATACGGAGGAAGCACCGATTATGACTCGAGCCAGTACAATGAAGACATGAAG  1764
      Y  S  S  Y  G  G  S  T  D  Y  D  S  S  Q  Y  N  E  D  M  K
1765 AAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTACAACGCCACGGGTGAGTACAGTAAT  1824
      K  F  R  K  M  A  L  G  T  Q  E  Y  N*  A  T  G  E  Y  S  N
1825 CCGACCAGTGACTATGGACTGTACCCGTCTGGTTCAAGCAGCGAGGGCCAAACCACACGC  1884
      P  T  S  D  Y  G  L  Y  P  S  G  S  S  S  E  G  Q  T  T  R
1885 GAAATGGAGATGGGGAAGATTAAGAGAACCGGTCAGGGTTATAGTGGACCTTCTCTTTAA  1944
      E  M  E  M  G  K  I  K  R  T  G  Q  G  Y  S  G  P  S  L  •
1945 accagatgggagagagaaattgaaggggtgtttttttcattattttttttaaaactgtaaagata  2004
2005 tgagaaaattgccttactctaattaaaaccactacgatataaggttataatacgttttga  2064
2065 attggttttttaaaaaaaaaaaaaaaaaaa  2093
```

Figure 1C

```
         10         20         30         40         50         60
TTAACTCTCT GGTCTCCGTG TCTCCTCTCT TCTCCTGCTG CTTCCTTTTA ACACTCTCTT
         70         80         90        100        110        120
CATTTGCCTT TTTGATTTAG ATCCAAAGAA GCAGAC_._T CCTCGGCGCC GTCTCCGGGG
        130        140        150        160        170        180
ACTGGTTCGC CTCCATCTCC ACCATCAAAC TCCACAACCA CCACTCCTCC TCCAGCTTCC
        190        200        210        220        230        240
GCTCCTCCTC CCACCACACC TTCTTCTCCT CCGCCGCCAT CCACTATTCC GACATCTCCT
        250        260        270        280        290        300
CCTCCTTCTT CTCGCTCTAC ACCTTCTGCT CCTCCTCCAT CTCCACCAAC TCCATCTACG
        310        320        330        340        350        360
CCGGGATCTC CACCTCCTCT TCCTCAGCCG TCTCCACCCG CTCCAACTAC GCCCGGATCT
        370        380        390        400        410        420
CCACCCGCAC CTGTTACTCC TCCTACTCGA AACCCTCCAC CTTCAGTCCC AGGACCACCG
        430        440        450        460        470        480
TCCAATCCTT CACGCGAAGG AGGATCTCCT CGACCTCCAT CTTCTCCCTC GCCGCCGTCT
        490        500        510        520        530        540
CCTTCTTCCG ACGGTTTATC AACAGGAGTG GTGGTGGGAA TCGCCATCGG AGGAGTCGCT
        550        560        570        580        590        600
CTGCTTGTGA TAGTGACTCT GATTTGTCTC CTCTGTAAGA AGAAACGACG GAGAGACGAA
        610        620        630        640        650        660
GAAGATGCTT ACTATGTTCC TCCGCCACCT CCTCCTGGTC CCAAAGCCGG AGGACCTTAC
        670        680        690        700        710        720
GGTGGACAGC AGCAACAATG GCGGCAACAA AACGCAACAC CACCGTCAGA TCATGTCGTG
        730        740        750        760        770        780
ACGTCACTAC CACCACCACC TAAGGCTCCA TCTCCACCAC GGCAACCTCC TCCACCTCCA
        790        800        810        820        830        840
CCACCGCCTT TCATGAGCAG CAGCGGCGGC TCCGACTACT CGGACCGTCC AGTTCTTCCT
        850        860        870        880        890        900
CCACCGTCTC CAGGGCTTGT GTTAGGCTTC TCCAAAAGCA CTTTCACATA CGAGGAGCTA
        910        920        930        940        950        960
GCTAGAGCCA CCAATGGTTT CTCCGAGGCG AACTTGTTAG GACAAGGCGG GTTCGGTTAC
        970        980        990       1000       1010       1020
GTGCACAAAG GTGTGTTGCC TAGTGGGAAA GAAGTTGCTG TGAAGCAGTT GAAAGTTGGG
       1030       1040       1050       1060       1070       1080
AGTGGTCAGG GAGAGAGGGA GTTTCAGGCA GAGGTTGAGA TCATCAGCAG AGTTCACCAC
       1090       1100       1110       1120       1130       1140
AGGCATCTGG TGTCTCTTGT TGGTTATTGC ATCGCCGGTG CCAAAAGATT GCTTGTCTAT
       1150       1160       1170       1180       1190       1200
GAGTTTGTTC CTAACAACAA TCTCGAGCTT CACCTCCATG GCGAGGACG GCCTACAATG
       1210       1220       1230       1240       1250       1260
GAATGGAGCA CCAGATTGAA GATTGCTCTT GGATCTGCTA AAGGACTTTC TTATCTTCAT
       1270       1280       1290       1300       1310       1320
GAAGATTGCA ATCCTAAAAT CATTCACCGT GATATCAAGG CTTCAAACAT ATTGATAGAT
       1330       1340       1350       1360       1370       1380
TTCAAGTTTG AAGCTAAGGT TGCTGATTTT GGTCTTGCTA AGATTGCTTC TGATACAAAC
       1390       1400       1410       1420       1430       1440
ACGCATGTAT CAACACGTGT GATGGGAACC TTTGGGTACT TGGCTCCGGA ATACGCTGCA
       1450       1460       1470       1480       1490       1500
AGCGGAAAGC TCACGGAGAA GTCTGACGTT TTCTCATTTG GCGTTGTGCT TTTGGAGCTC
       1510       1520       1530       1540       1550       1560
```
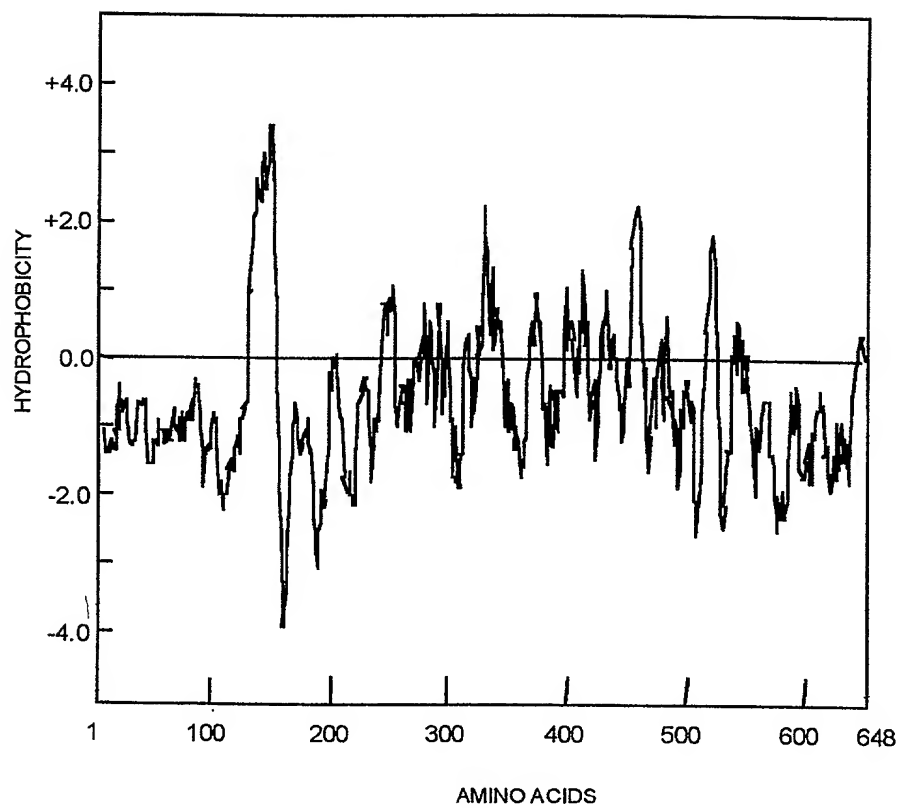
Figure 1(d)   (continued on next page)

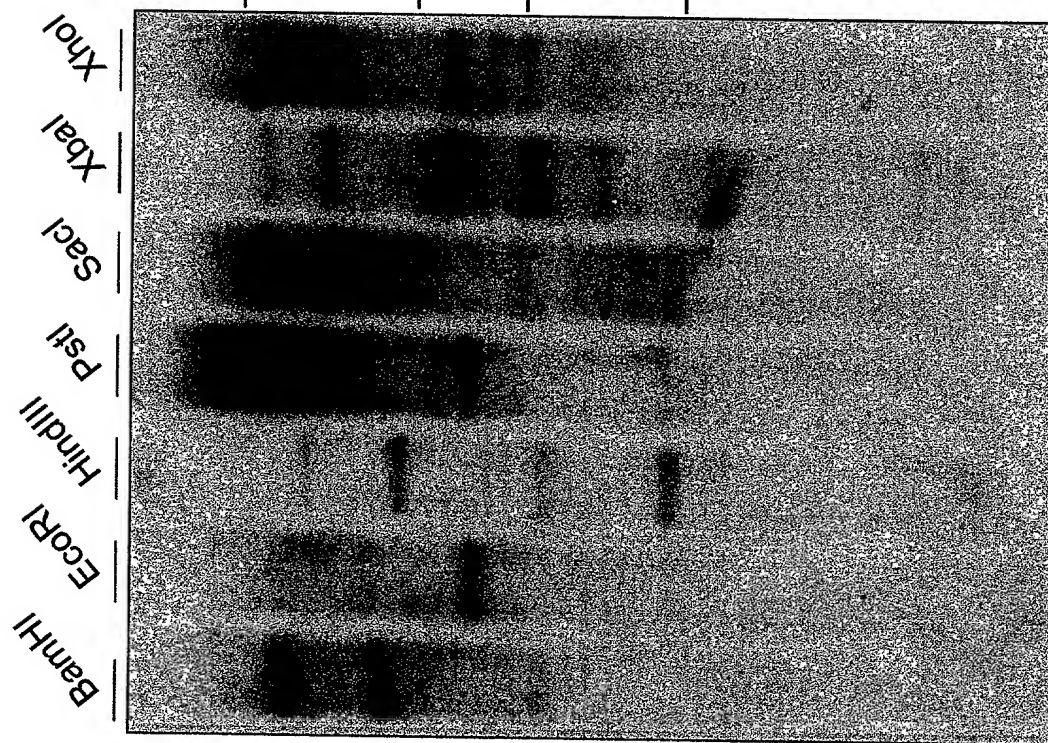```
ATTACTGGAC GTCGACCCGT TGATGCCAAC AATGTCTATG TAGATGACAG CTTAGTTGAC
     1570   1580   1590   1600   1610   1620
TGGGCACGAC CATTGCTTAA CCGAGCATCT GAGCAAGGAG ACTTTGAGGG TTTAGCTGAT
     1630   1640   1650   1660   1670   1680
GCAAAGATGA ATAATGGGTA TGACAGAGAG GAGATGGCTC GCATGGTTGC TTGTGCTGCG
     1690   1700   1710   1720   1730   1740
GCTTGTGTTC GCCATTCAGC TCGCCGCAGA CCTCGCATGA GCCAGATTGT GCGTGCGTTA
     1750   1760   1770   1780   1790   1800
GAAGGAAATG TATCACTGTC AGATCTTAAC GAAGGGATGA GACCAGGTCA AAGCAATGTA
     1810   1820   1830   1840   1850   1860
TACAGCTCAT ACGGAGGAAG CACCGATTAT GACTCGAGCC AGTACAATGA AGACATGAAG
     1870   1880   1890   1900   1910   1920
AAGTTTAGGA AAATGGCACT TGGAACTCAA GAGTACAACG CCACGGGTGA GTACAGTAAT
     1930   1940   1950   1960   1970   1980
CCGACCAGTG ACTATGGACT GTACCCGTCT GGTTCAAGCA GCGAGGGCCA AACCACACGC
     1990   2000   2010   2020   2030   2040
GAAATGGAGA TGGGGAAGAT TAAGAGAACC GGTCAGGGTT ATAGTGGACC TTCTCTTTAA
     2050   2060   2070   2080   2090   2100
ACCAGATGGG AGAGAAATTG AAGGGTGTTT TTTCATTATT TTTTTAAAAC TGTAAAGATA
     2110   2120   2130   2140   2150   2160
TGAGAAAATT GCCTTACTCT AATTAAAACC ACTACGATAT AAGGTTATAA TACGTTTTGA
     2170   2180   2190   2200   2210   2220
GATTGGTTTTT AAAAAAAAAA AAAAAAAAA. ....
```

Figure 1(d) (continued)

```
LTLWSPCLLS SPAASF*HSL HLPF*FRSKE AD·ISSAPSPG TGSPPSPPSN STTTTPPPAS
APPPTTPSSP PPPSTIPTSP PPSSRSTPSA PPPSPPTPST PGSPPPLPQP SPPAPTTPGS
PPAPVTPPTR NPPPSVPGPP SNPSREGGSP RPPSSPSPPS PSSDGLSTGV VVGIAIGGVA
LLVIVTLICL LCKKKRRRDE EDAYYVPPPP PPGPKAGGPY GGQQQQWRQQ NATPPSDHVV
TSLPPPPKAP SPPRQPPPPP PPPFMSSSGG SDYSDRPVLP PPSPGLVLGF SKSTFTYEEL
ARATNGFSEA NLLGQGGFGY VHKGVLPSGK EVAVKQLKVG SGQGEREFQA EVEIISRVHH
RHLVSLVGYC IAGAKRLLVY EFVPNNNLEL HLHGEGRPTM EWSTRLKIAL GSAKGLSYLH
EDCNPKIIHR DIKASNILID FKFEAKVADF GLAKIASDTN THVSTRVMGT FGYLAPEYAA
SGKLTEKSDV FSFGVVLLEL ITGRRPVDAN NVYVDDSLVD WARPLLNRAS EQGDFEGLAD
AKMNNGYDRE EMARMVACAA ACVRHSARRR PRMSQIVRAL EGNVSLSDLN EGMRPGQSNV
YSSYGGSTDY DSSQYNEDMK KFRKMALGTQ EYNATGEYSN PTSDYGLYPS GSSSEGQTTR
EMEMGKIKRT GQGYSGPSL   TRWERN*RVF FHYFFKTVKI *ENCLTLIKT TTI*GYNTF*
IGF*KKKKK
```
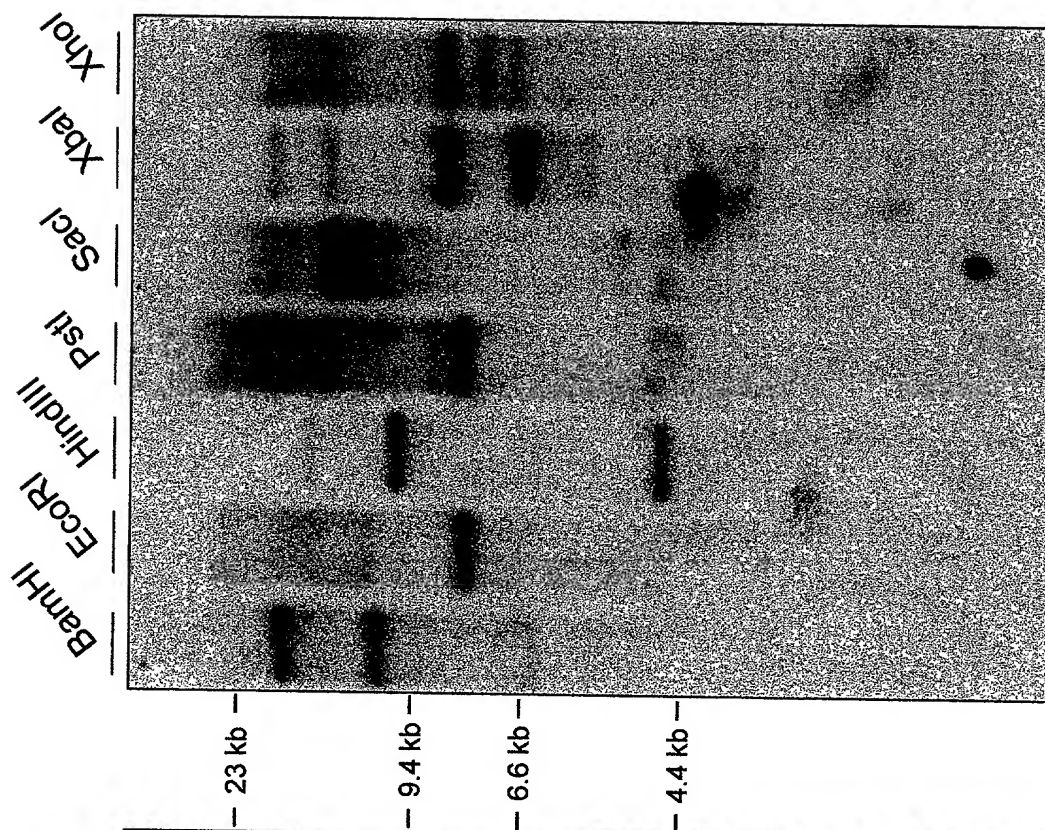
Figure 1(e)

Figure 2

Figure 3

**A.**



Figure 4A

Figure 4B

min                hr

0   5   15  30  45   1   2   4

PERK1
2.2 kb

Cyclophilin
700 bp
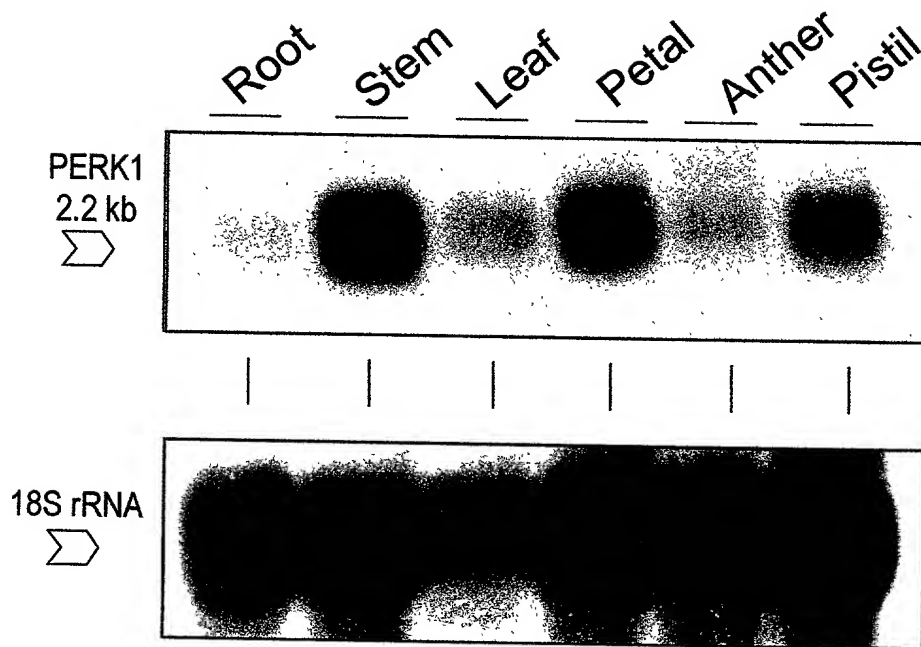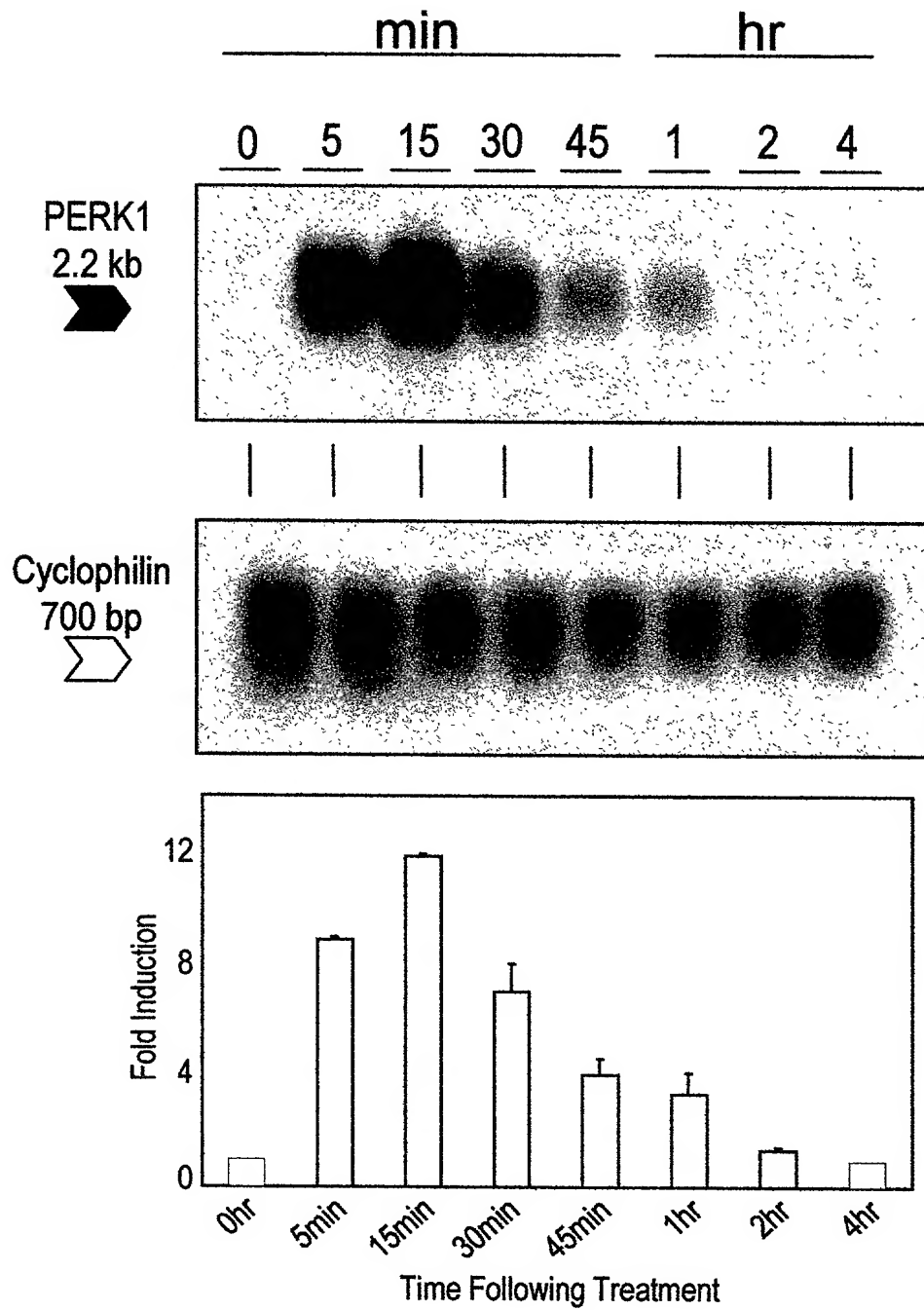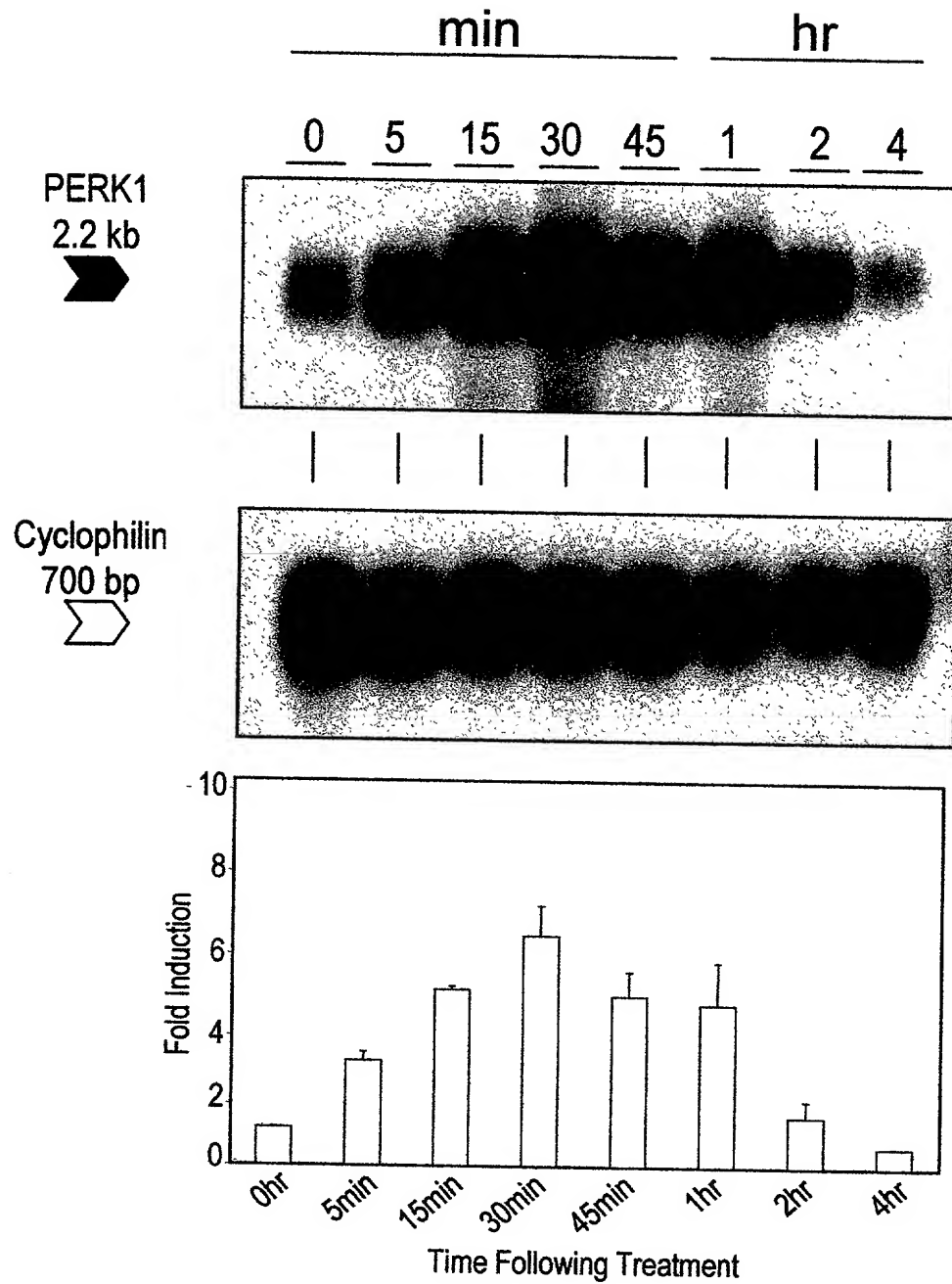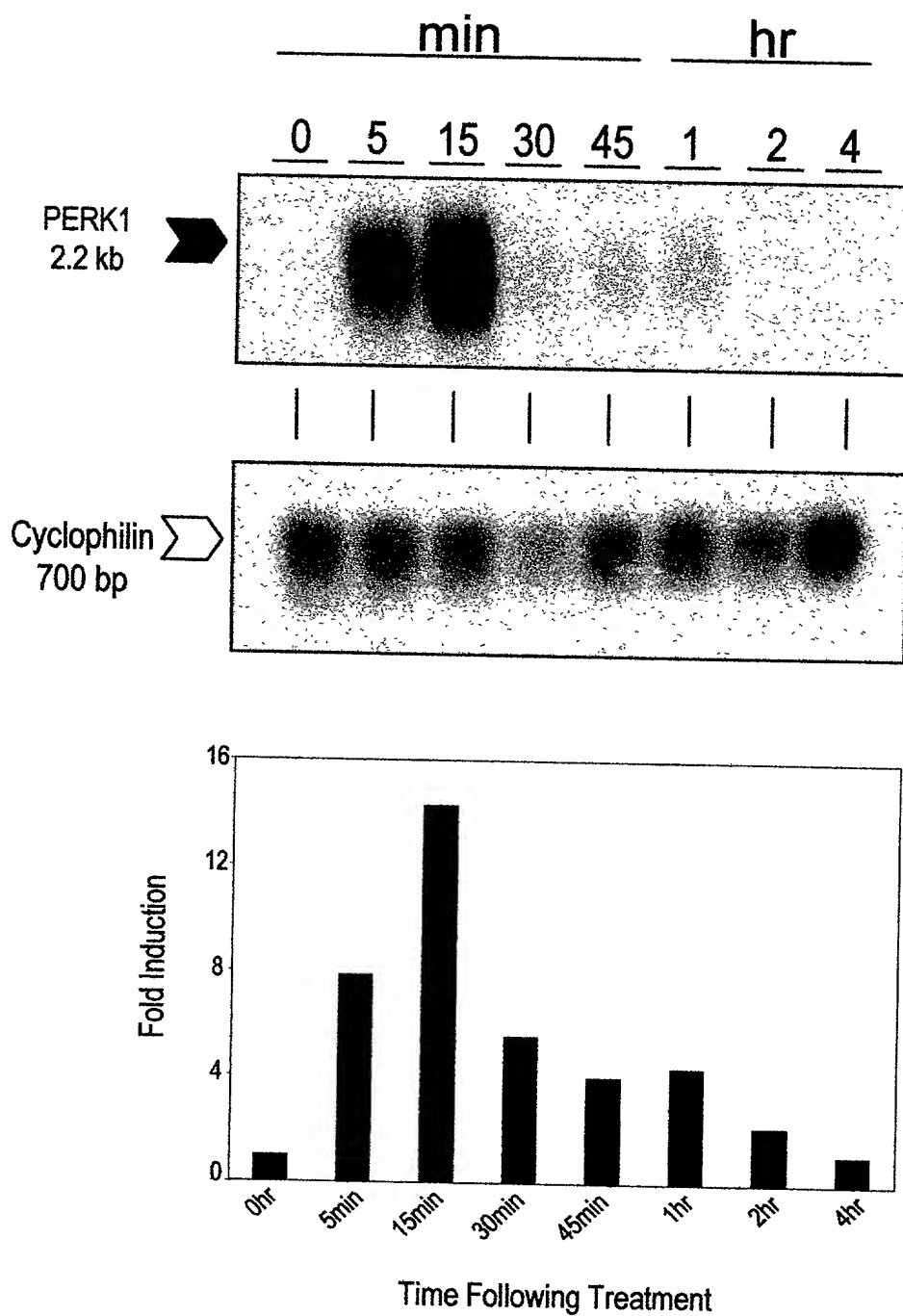
Figure 4a

# A.
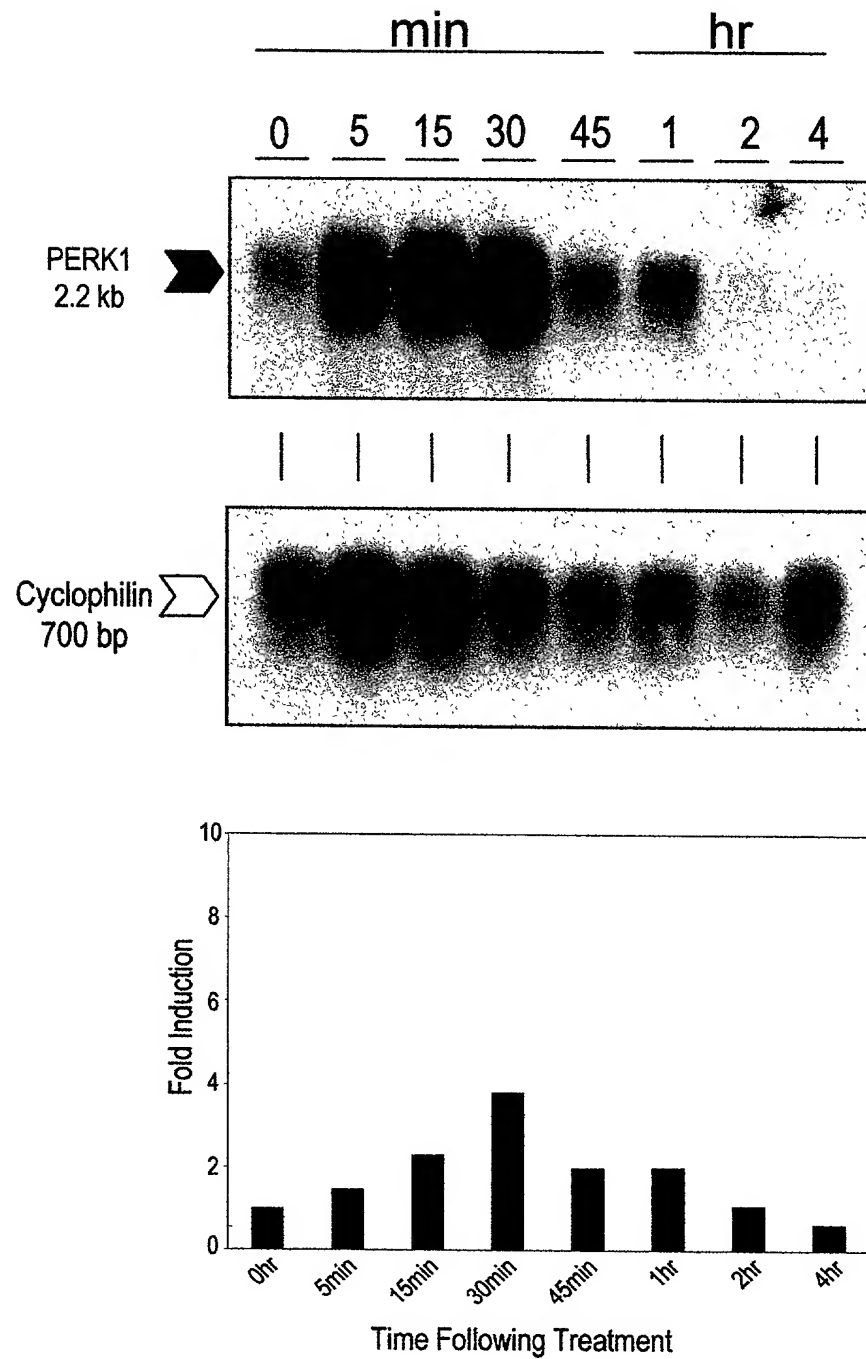


Figure 4b(A)

# B.



Figure 4b(B)

Figure 4c

**A.**

min | hr

0 5 15 30 45 | 1 2 4 12 24 36 48

◁ PERK1 2.2kb

◀ Cyclophilin 700bp

10.0

8.0

6.0

Fold Induction

4.0

2.0

0.0

0hr 5min 15min 30min 45min 1hr 2hr 4hr 12hr 24hr 36hr 48hr

Time Following Treatment
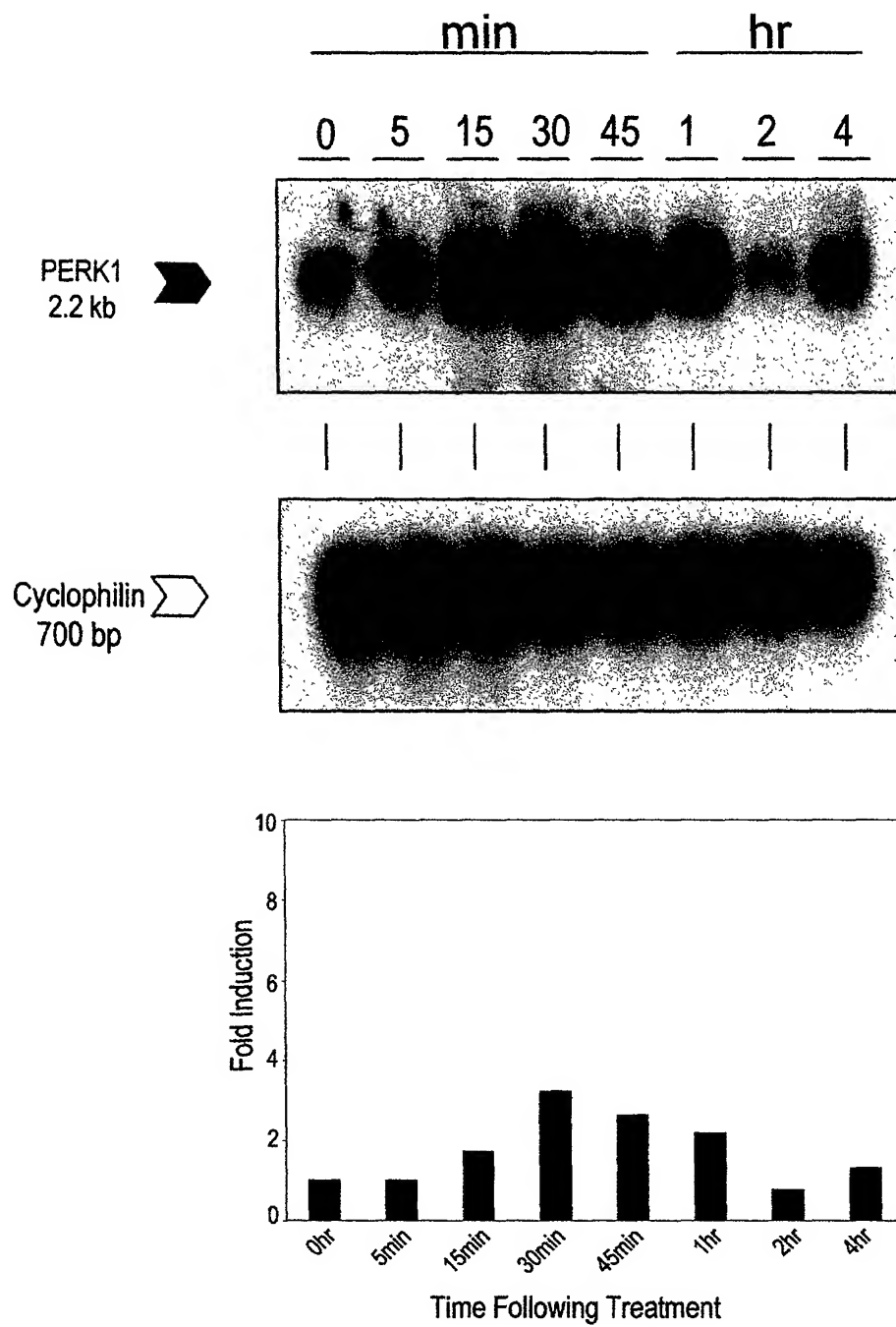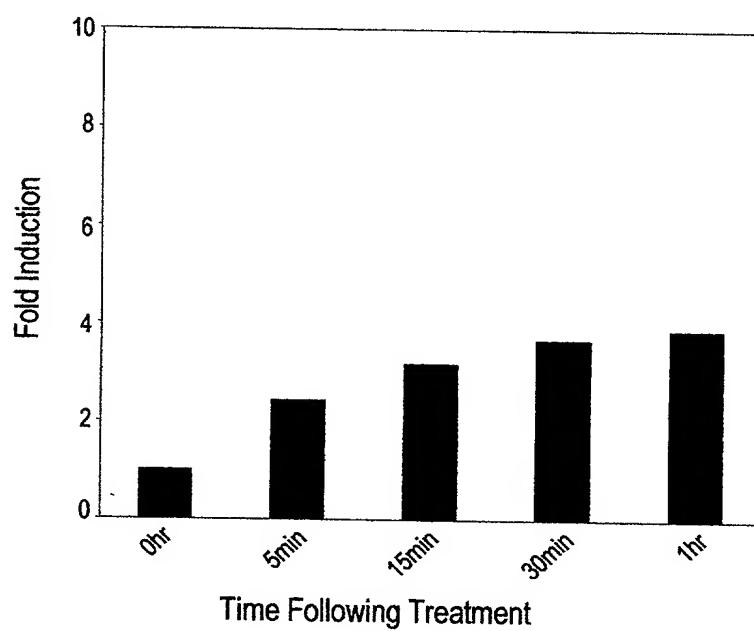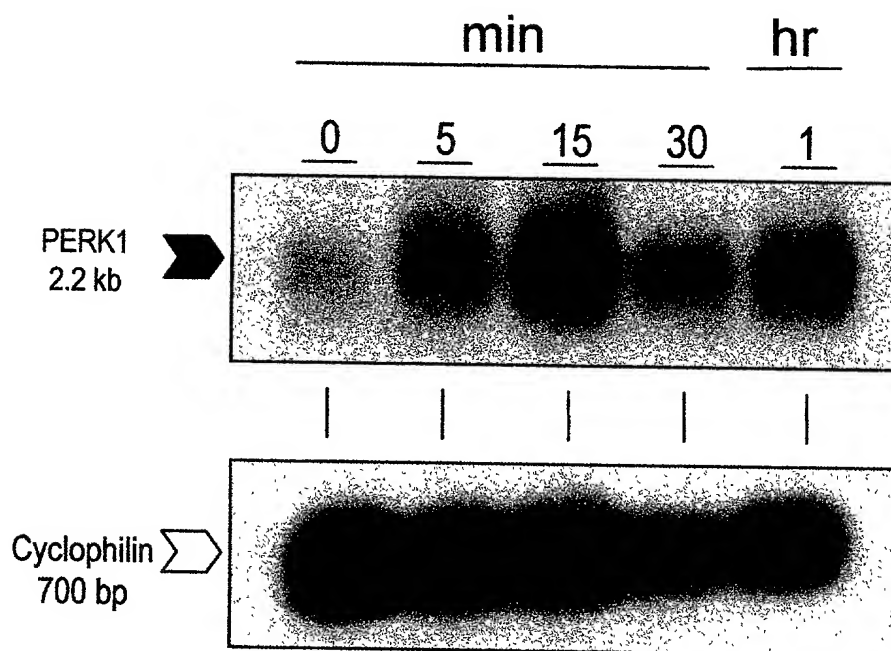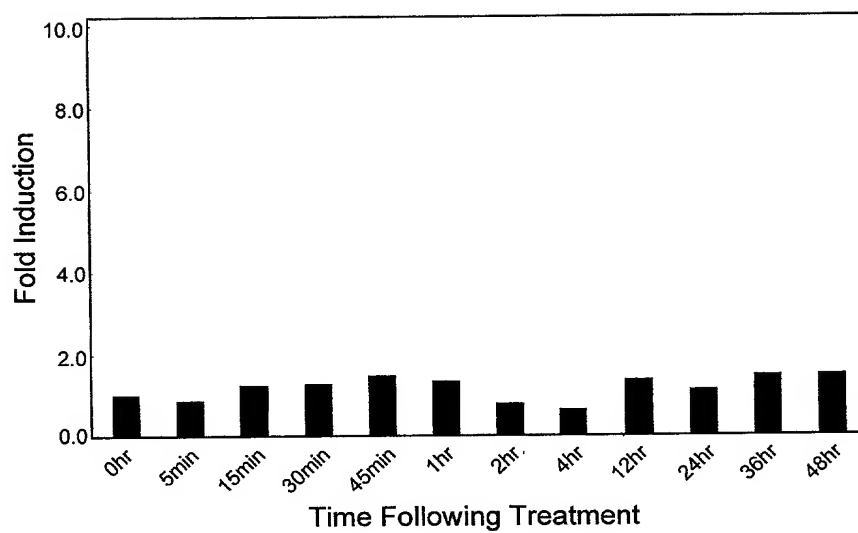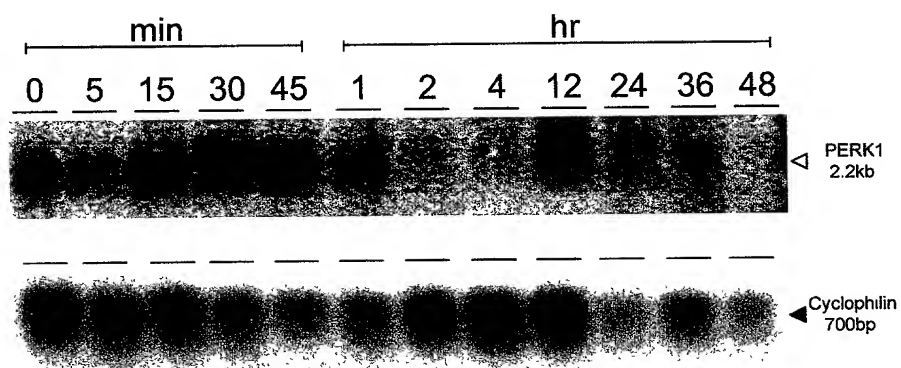
Figure 5A

**B.**
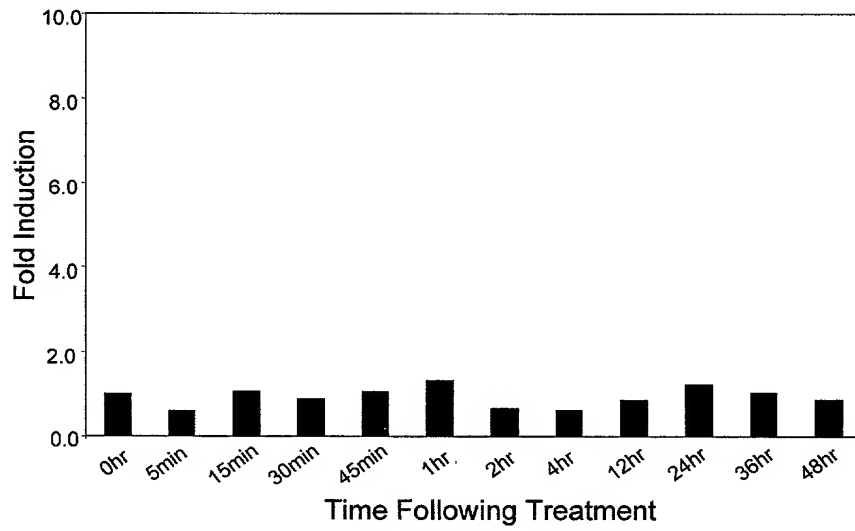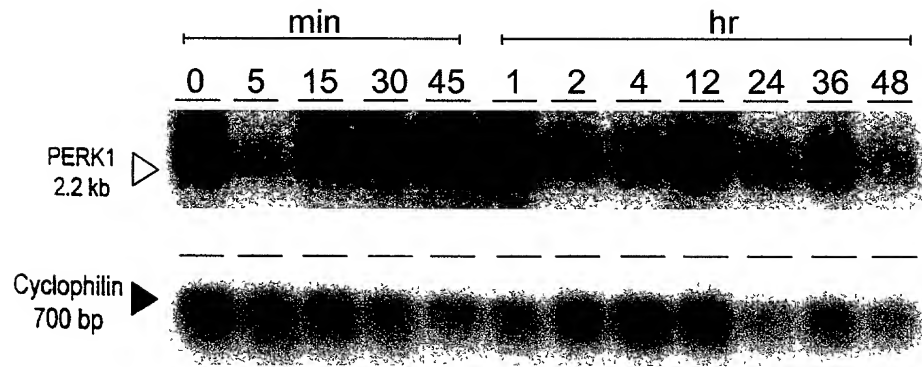


Figure 5B

# A.
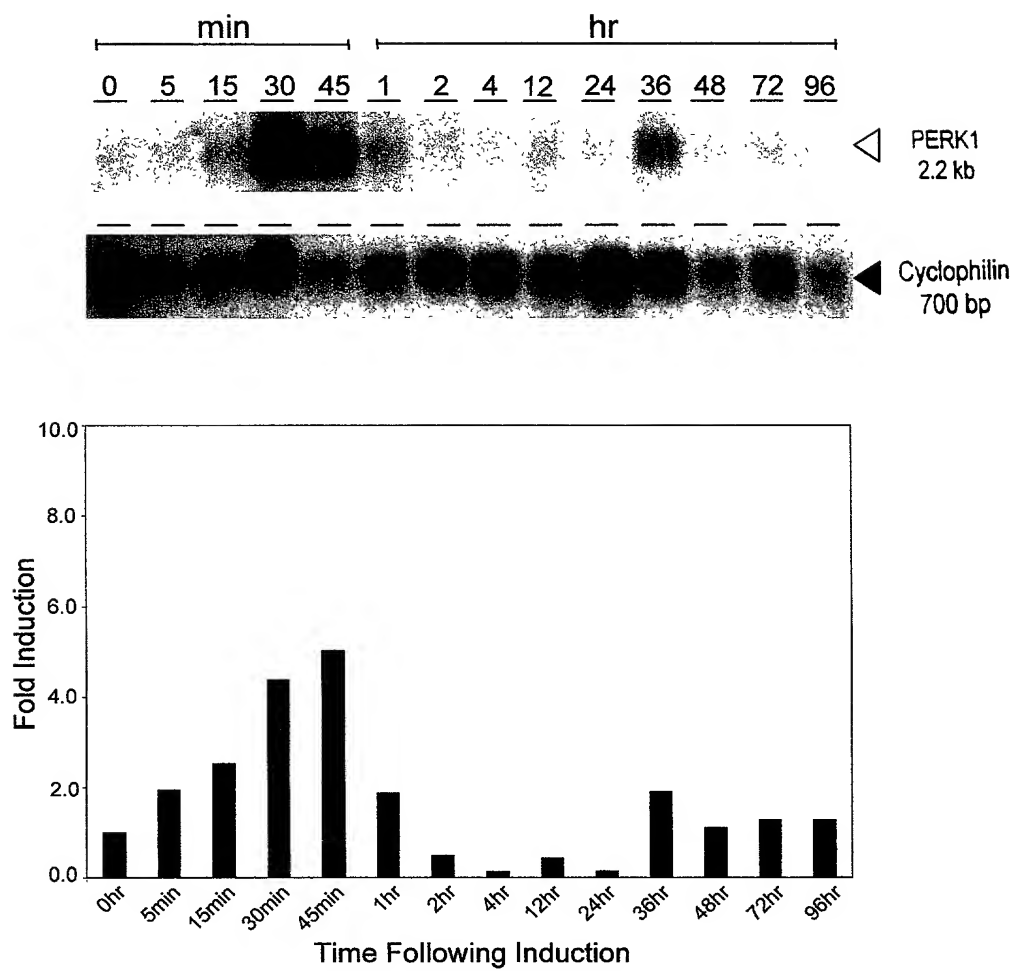


Figure 6A

**B.**
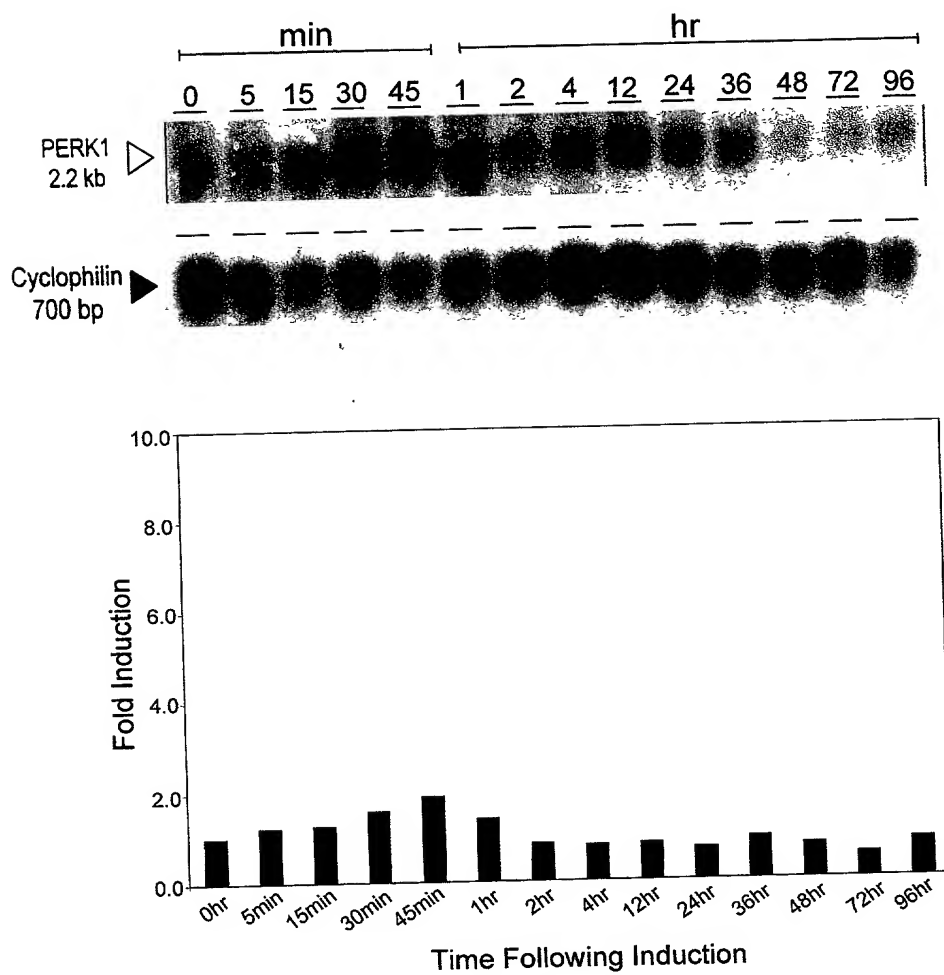


Figure 6B

# • Fungal pathogen : *Sclerotinia sclerotiorum*

hrs: 2 3 4 5 6 6.5 7 7.5 8 8.5 9 9.5 10 10.5 11 11.5 12 12.5 13 13.5 14 14.5 15 15.5 16 16.5

PERK1

Cyclophilin

**TREATMENT**

\* Poor adherence of fungal agar plugs

**CONTROL**

Fold Induction

Time Following Treatment (hr)
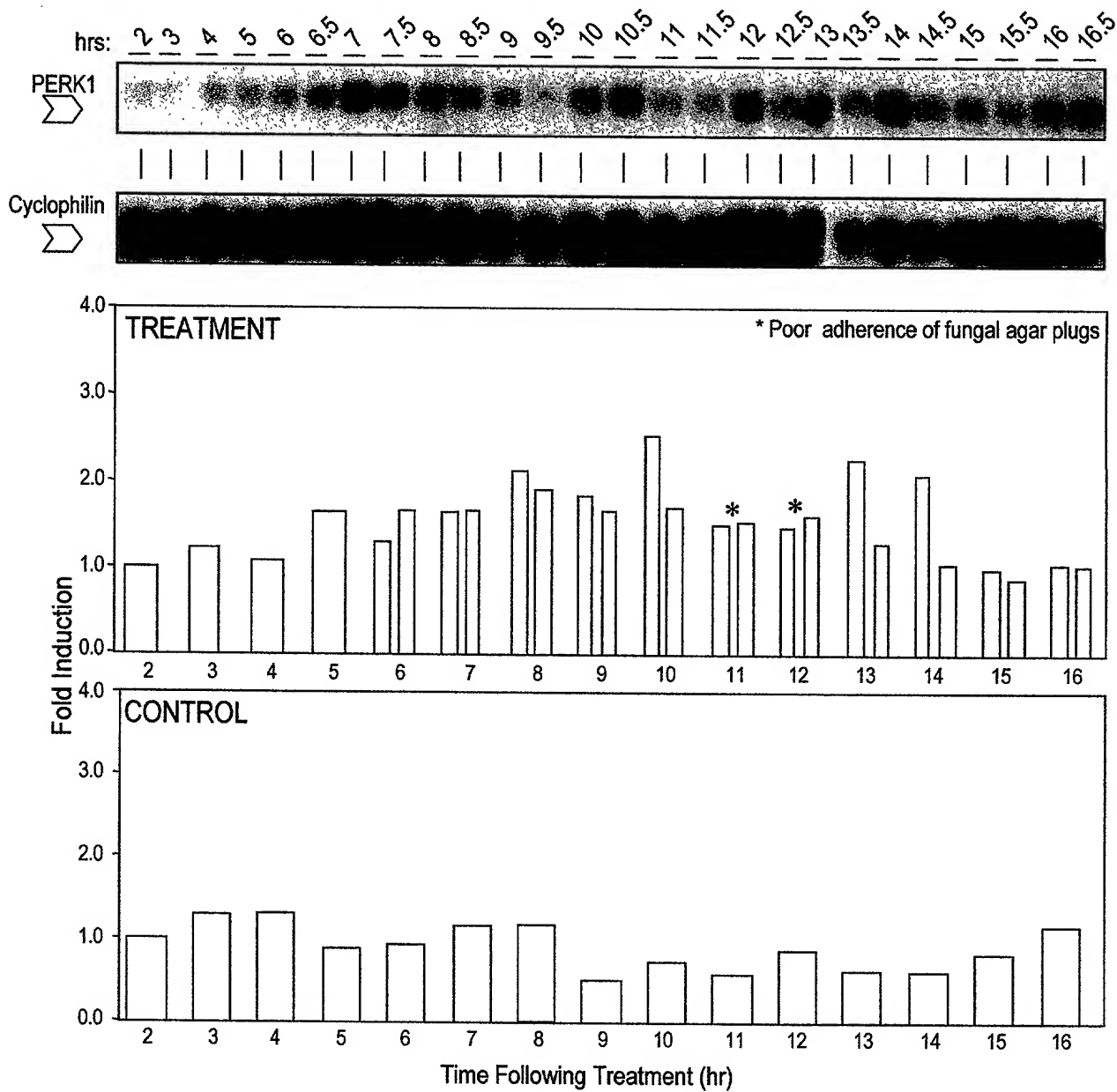
Figure 7

Mechanical Wounding
or
Insect Attack

Pathogen
Attack
(eg. bacterial, fungal)

MeJA
Independent
Pathway

MeJA
Dependent
Pathway

?

SA

PERK1

PERK1

PERK1

Figure 8

# A.

Purified CDmut_{K-E}    Purified CDwt

175kDa —

83kDa —

62kDa —

47.5kDa —

Lanes:        1            2

# B.

Purified CDmut_{K-E}    Purified CDwt

175kDa —

83kDa —

62kDa —

Lanes:        1            2

# C.

CDwt    CDmut_{K-E}

175kDa —

83kDa —

62kDa —

Lanes:        1            2

# D.

pS

pT

pY

Figure 9

PERK1

†AAC98010
36% (76-211)    30%    50% (330-704)

† AAD15491
36% (5-134)    30%    64% (173-614)

† CAA18823
36% (1-133)    33%    52% (214-674)

† CAA18590
32% (50-185)    33%    63% (370-710)
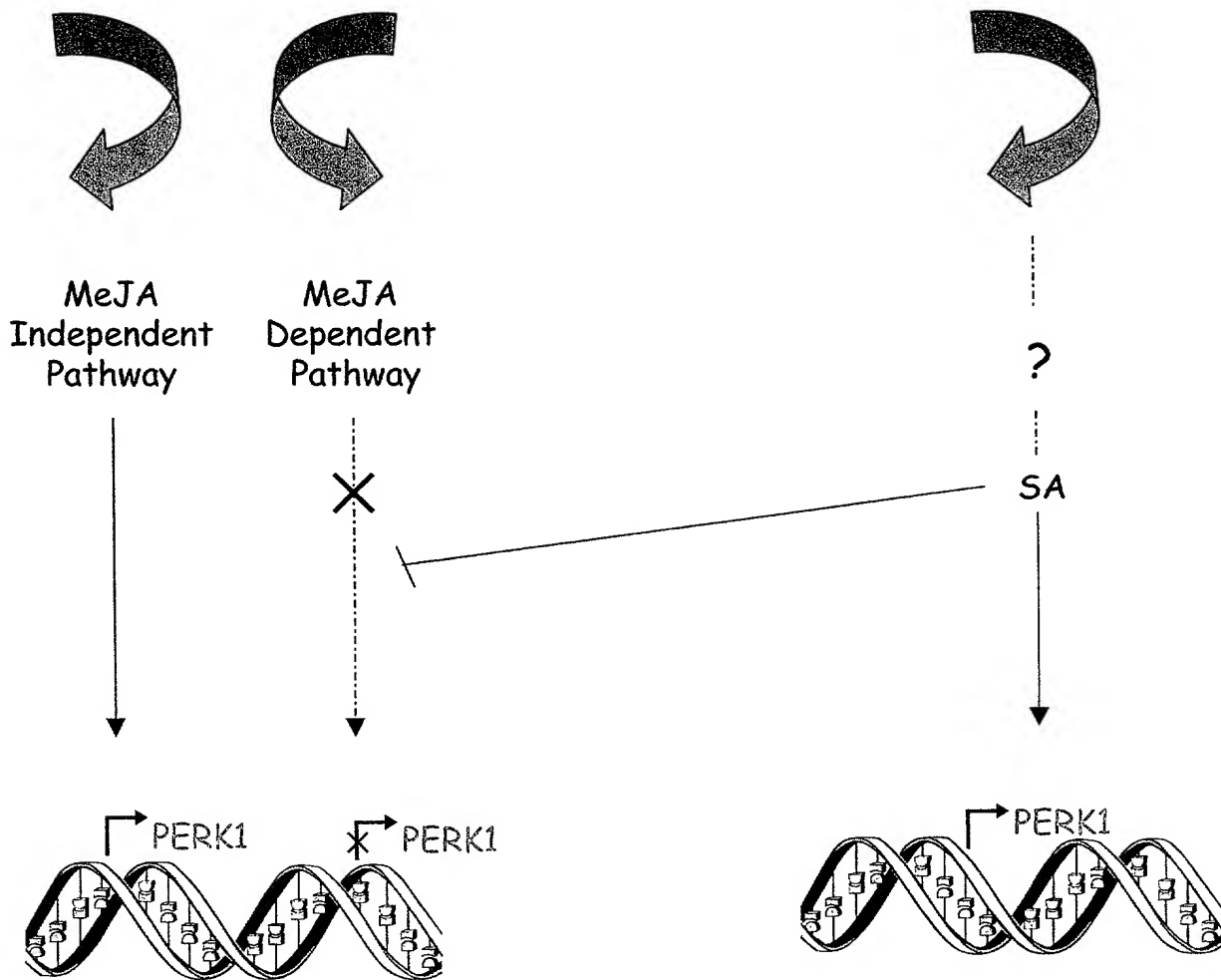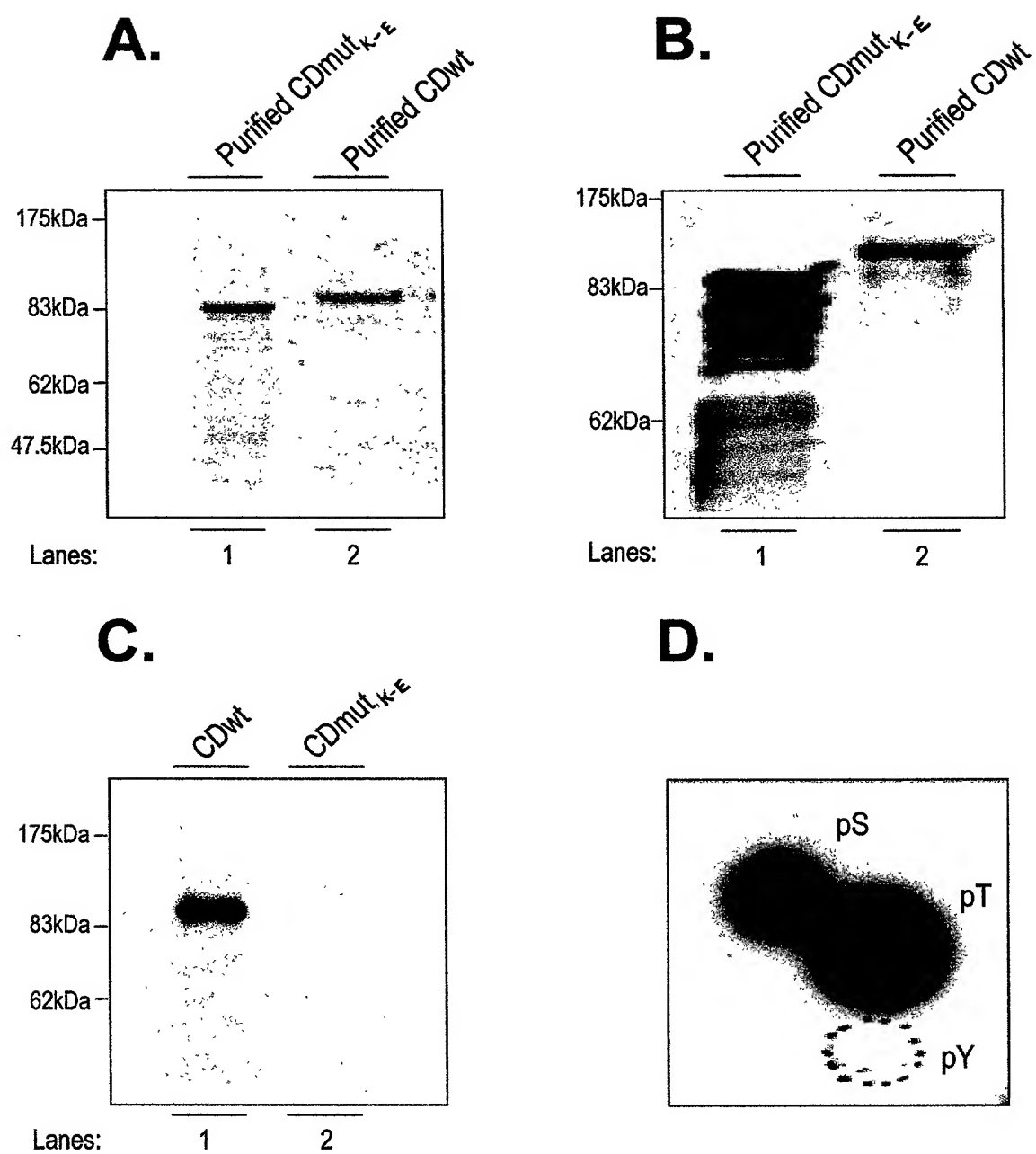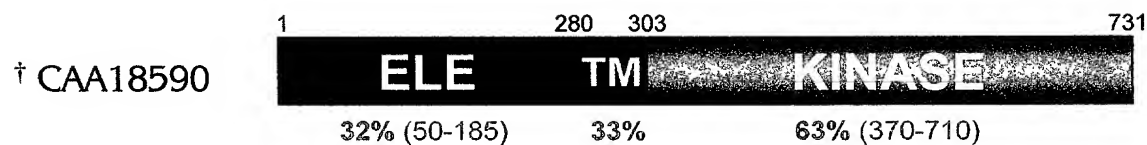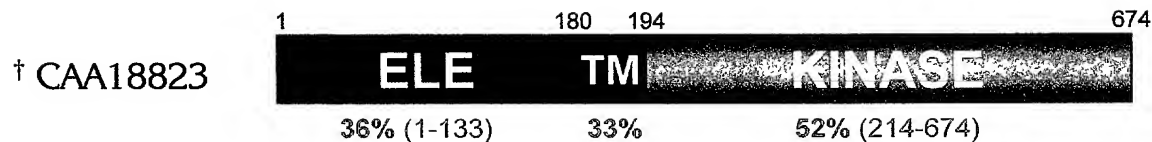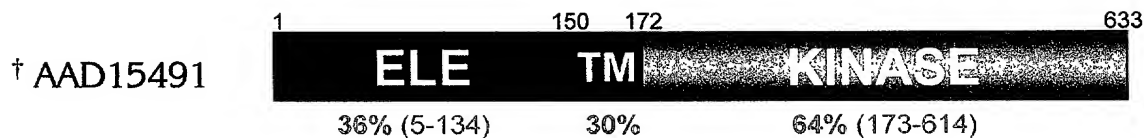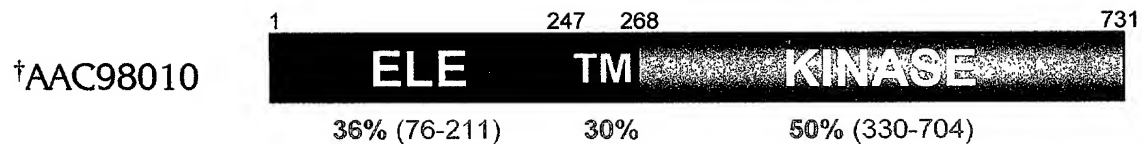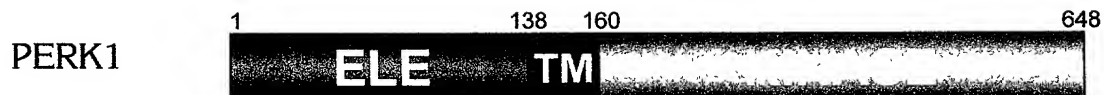
All have no signal peptide and predicted
to be Type1b integral membrane proteins

† Predicted proteins from the Arabidopsis
genome sequencing project.

ELE: Extensin-like extracellular ;  TM: Transmembrane domain ;  Red = sequence identity to PERK1
domain

| | PERK1 | ACC98010 | AAD15491 | CAA18823 | CAA18590 | |
|---|---|---|---|---|---|---|
| PERK1 | | 50% | 64% | 52% | 63% | |
| ACC98010 | 36% | | 54% | 50% | 52% | Kinase |
| AAD15491 | 36% | 25% | | 63% | 51% | Domains |
| CAA18823 | 36% | 29% | 27% | | 60% | |
| CAA18590 | 32% | 36% | 30% | 27% | | |

ELE Domains

Figure 10

TAGAAAAAAA AAAATGTCAG ACTTAGGCGA GTCGCCGAGT TCTTCACCAC CAGCACCACC
AGCTGATACC GCTCCTCCAC CAGAGACTCC ATCAGAAAAC TCAGCTCTTC CACCTGTTGA
TTCCTCTCCT CCTAGTCCAC CAGCTGATTC ATCATCAACA CCGCCGCTGT CAGAACCATC
CACTCCTCCT CCAGATTCAC AGCTTCCTCC TTTACCTTCG ATTCTTCCTC CGCTAACAGA
TTCTCCACCT CCACCTTCCG ATTCTTCTCC ACCCGTTGAT TCAACCCCTT CTCCGCCGCC
ACCGACGTCA AACGAATCTC CTTCTCCTCC AGAAGATTCC GAAACACCAC CTGCTCCACC
AAATGAATCC AATGACAACA ACCCTCCTCC GTCTCAAGAT CTTCAATCGC CTCCTCCATC
GTCGCCGTCG CCGAATGTAG GACCCACAAA CCCGGAATCA CCACCGTTAC AATCTCCTCC
AGCTCCACCA GCATCAGATC CTACAAATTC ACCGCCAGCT TCACCATTAG ACCCTACCAA
TCCTCCCCCA ATACAACCAT CAGGACCAGC CACTTCTCCT CCGGCTAATC CCAACGCTCC
GCCGAGCCCA TTCCCCACAG TACCACCCAA AACTCCTTCT AGTGGACCTG TGGTGTCTCC
ATCTCTCACA TCCCCTAGTA AAGGAACTCC TACTCCAAAC CAAGGCAATG GAGATGGCGG
TGGCGGTGGT GGCGGCTATC AAGGGAAGAC TATGGTTGGT ATGGCTGTAG CCGGTTTCGC
AATCATGGCG CTTATAGGCG TTGTGTTCTT AGTGAGAAGA AGAAAAAGA GAAACATTGA
TAGCTATAAT CACTCACAGT ACTTGCCACA TCCCAATTTC TCTGTTAAAT CAGGTTTAAA
AATCTCACCT TTATCTCTCT CTGATCATCT TCTATGTGCT TGAATCATCT CTCTGACTAT
CTTTGCTTTT GATGTAGATG GATTCTTATA CGGTCAAGAT CCAGGTAAAG GATACTCCTC
TGGTCCTAAT GGTTCAATGT ATAACAATTC ACAGCAACAA CAATCCTCTA TGGGAAACAG
TTATGGTACA GCTGGTGGTG GTTATCCTCA TCATCAAATG CAATCAAGTG GCACACCTGA
CTCTGCTATA CTCGGAAGTG GCCAGACTCA TTTCAGTTAC GAAGAGCTTG CTGAGATAAC
ACAAGGCTTT GCTCGCAAAA ACATTCTTGG AGAAGGCGGA TTTGGATGTG TCTATAAAGG
TACATTGCAG GATGGTAAAG TTGTTGCGGT TAAGCAGCTT AAAGCTGGAA GTGGACAAGG
TGACCGTGAA TTCAAAGCAG AGGTTGAGAT CATCAGCCGC GTTCATCATC GCCATTGGT
CTCTCTGGTT GGTTACTGCA TTTCAGACCA GCATAGATTG CTTATCTATG AGTATGTTTC
TAATCAAACC TTGGAGCATC ATTTGCATGG TGAGTGACTT GTTACCATTT TCGTTATAGA
TAAGACTTTT TTTTAGCTTT ACGTGTTAGA CTGACTCGCT TTACGCTTTA GGAAAGGGTT
TGCCAGTTTT AGAGTGGTCT AAGAGAGTCC GGATCGCTAT AGGATCAGCC AAAGGGTTGG
CATATCTTCA CGAAGACTGT AAGTAATGCC TTCACATTTT CTTAGTTGTG TGCTTTGGTT
ATGCACTTCA TAGTTTAAAC AGAAGCCAAA AATCATATCC TTGTTTTATT TTACAGGTCA
TCCGAAAATC ATTCACAGAG ATATAAAGTC AGCAAATATT CTTCTAGATG ATGAATATGA
AGCTCAGGCA ATAATGAAAT CCTCCTTTTC GTTAAATCTA TCTTATGACT GTAAAGTTTT
AGTTAATGAG ACTTGTTCTG TTTTTTGGAT GTTTAGGTTG CTGATTTTGG ACTTGCTAGA
CTCAATGATA CAACACAAAC TCATGTTTCA ACTCGGGTTA TGGGAACCTT CGGGTAAGCA
AACATTCATC ACAAACTCTA CTCCAAAACT GGACCTTATT GATCCAATGC CTGATGAAAA
GTTTGTTATA TATGGCTTGA GGCAACAAAT TGGATCAAAC CTGAATCTTT ATTGATCGTA
TGGCTGCATG ACATGTTTTG TGTTAAGGTA CCTAGCGCCG GAATATGCAT CAAGTGGAAA
ATTGACTGAT AGATCCGATG TATTCTCATT CGGGGTTGTT CTCTTAGAGC TTGTAACTGG
ACGGAAACCA GTTGACCAGA CTCAGCCTCT AGGAGAAGAG AGTTTGGTTG AATGGGTAAG
AATCCAACTT TCAAACATTC TTCAATAATA GTAAGATTGG CCCTAGTATA CTTATATAGT
ACTTATAAAT GAACTCACAG GCCGCGCCCGC TGCTTCTCAA AGCCATTGAG ACCGGAGATT
TAAGCGAACT GATTGATACA CGGCTTGAAA AGCGTTATGT GGAGCATGAA GTCTTCAGAA
TGATCGAGAC AGCCGCTGCA TGTGTTAGAC ATTCTGGTCC AAAACGCCCA CGCATGGTTC
AGGTAATTCT GACTAACCAA AAGTCCAAAG CTCCCATATA TAGTAACAAG TGATTTCTCA
CATCTGAAAA CTTATCTACT CTTCGAAATA AGGTTGTGAG AGCATTGGAC TGCGACGGAG
ACTCGGGAGA TATTAGCAAC GGAATCAAAA TTGGGCAAAG CACAACTTAT GACTCAGGGC
AATACAATGA AGACATTATG AAATTCAGGA AAATGGCGTT TGGTGGTGAT AACAGCGTAG
AGTCAGGATT GTACAGTGGA AACTACTCTG CCAAAAGCTC TTCAGATTTC TCAGGGAATG
AATCTGAGAC TCGGCCTTTC AACAACCGAC GGTTCTGATC ATACAATAGG TGAAAGTAAC

```
   1    M   S   D   L   G   E   S   P   S   S   S   P   P   A   P   P   A   D   T   A     20
   1  ATG TCA GAC TTA GGC GAG TCG CCG AGT TCT TCA CCA CCA GCA CCA CCA GCT GAT ACC GCT     60
  21    P   P   P   E   T   P   S   E   N   S   A   L   P   P   V   D   S   S   P   P     40
  61  CCT CCA CCA GAG ACT CCA TCA GAA AAC TCA GCT CTT CCA CCT GTT GAT TCC TCT CCT CCT    120
  41    S   P   P   A   D   S   S   S   T   P   P   L   S   E   P   S   T   P   P   P     60
 121  AGT CCA CCA GCT GAT TCA TCA ACA CCG CCG CTG TCA GAA TCC ACT CCT CCT CCA           180
  61    D   S   Q   L   P   P   L   P   S   I   L   P   P   L   T   D   S   P   P   P     80
 181  GAT TCA CAG CTT CCT CCT TTA CCT TCG ATT CTT CCT CCG CTA ACA GAT TCT CCA CCT CCA    240
  81    P   S   D   S   S   P   P   V   D   S   T   P   S   P   P   P   P   T   S   N    100
 241  CCT TCC GAT TCT TCT CCA CCC GTT GAT TCA ACC CCT TCT CCG CCG CCA ACG TCA AAC        300
 101    E   S   P   S   P   P   E   D   S   E   T   P   P   A   P   P   N   E   S   N    120
 301  GAA TCT CCT TCT CCT CCA GAA GAT TCC GAA ACA CCA CCT GCT CCA CCA AAT GAA TCC AAT    360
 121    D   N   N   P   P   P   S   Q   D   L   Q   S   P   P   P   S   S   P   S   P    140
 361  GAC AAC AAC CCT CCT CCG TCT CAA GAT CTT CAA TCG CCT CCA TCG TCG CCG TCG CCG        420
 141    N   V   G   P   T   N   P   E   S   P   P   L   Q   S   P   P   A   P   P   A    160
 421  AAT GTA GGA CCC ACA AAC CCG GAA TCA CCA CCG TTA CAA TCT CCT CCA GCT CCA CCA GCA    480
 161    S   D   P   T   N   S   P   P   A   S   P   L   D   P   T   N   P   P   P   I    180
 481  TCA GAT CCT ACA AAT TCA CCG CCA GCT TCA CCA TTA GAC CCT AAC CCT CCC CCA ATA        540
 181    Q   P   S   G   P   A   T   S   P   P   A   N   P   N   A   P   P   S   P   F    200
 541  CAA CCA TCA GGA CCA GCC ACT TCT CCT CCG GCT AAT CCC AAC GCT CCG CCG AGC CCA TTC    600
 201    P   T   V   P   P   K   T   P   S   S   G   P   V   V   S   P   S   L   T   S    220
 601  CCC ACA GTA CCA CCA AAA ACT CCT TCT AGT GGA CCT GTG GTG GGA TCT CTC ACA TCC        660
 221    P   S   K   G   T   P   T   P   N   Q   G   N   G   D   G   G   G   G   G   G    240
 661  CCT AGT AAA GGA ACT CCT ACT CCA AAC CAA GGC AAT GGA GAT GGC GGT GGC GGT GGT GGC    720
 241    G   Y   Q   G   K   T  _M   V   G   M   A   V   A   G   F   A   I   M   A   L_   260
 721  GGC TAT CAA GGG AAG ACT ATG GTT GGT ATG GCT GTA GCC GGT TTC GCA ATC ATG GCG CTT    780
 261   _I   G   V   V   F   L   V_   R   R   K   K   K   R   N   I   D   S   Y   N   H    280
 781  ATA GGC GTT GTG TTC TTA GTG AGA AGA AAG AAA AAG AGA AAC ATT GAT AGC TAT AAT CAC    840
 281    S   Q   Y   L   P   H   P   N   F   S   V   K   S   D   G   F   L   Y   G   Q    300
 841  TCA CAG TAC TTG CCA CAT CCC AAT TTC TCT GTT AAA TCA GAT GGA TTC TTA TAC GGT CAA    900
 301    D   P   G   K   G   Y   S   S   G   P   N   G   S   M   Y   N   N   S   Q   Q    320
 901  GAT CCA GGT AAA GGA TAC TCC TCT GGT CCT AAT GGT TCA ATG TAT AAC AAT TCA CAG CAA    960
 321    Q   Q   S   S   M   G   N   S   Y   G   T   A   G   G   G   Y   P   H   H   Q    340
 961  CAA CAA TCC TCT ATG GGA AAC AGT TAT GGT ACA GCT GGT GGT GGT TAT CCT CAT CAT CAA   1020
 341    M   Q   S   S   G   T   P   D   S   A   I   L   G   S   G   Q   T   H   F   S    360
1021  ATG CAA TCA AGT GGC ACA CCT GAC TCT GCT ATA CTC GGA AGT GGC CAG ACT CAT TTC AGT   1080
 361    Y   E   E   L   A   E   I   T   Q   G   F   A   R   K   N   I   L   E   G        380
1081  TAC GAA GAG CTT GCT GGA ATA ACA CAA GGC TTT GCT CGC AAA AAC ATT CTT GGA GAA GGC   1140
 381    G   F   G   C   V   Y   K   G   T   L   Q   D   G   K   V   V   A   V   K   Q    400
1141  GGA TTT GGA TGT GTC TAT AAA GGT ACA TTG CAG GAT GGT AAA GTT GTT GCG GTT AAG CAG   1200
 401    L   K   A   G   S   G   Q   G   D   R   E   F   K   A   E   V   E   I   I   S    420
1201  CTT AAA GCT GGA AGT GGA CAA GGT GAC CGT GAA TTT AAA GCA GAG GTT GAG ATC ATC AGC   1260
 421    R   V   H   H   R   H   L   V   S   L   V   G   Y   C   I   S   D   Q   H   R    440
1261  CGC GTT CAT CAT CGC CAT TTG GTC TCT CTG GTT GGT TAC TGC ATT TCA GAC CAG CAT AGA   1320
 441    L   L   I   Y   E   Y   V   S   N   Q   T   L   E   H   H   L   H   E   W   S    460
1321  TTG CTT ATC TAT GAG TAT GTT TCT AAT CAA ACC TTG GAG CAT CAT TTG CAT GAG TGG TCT   1380
 461    K   R   V   R   I   A   I   G   S   A   K   G   L   A   Y   L   H   E   D   C    480
1381  AAG AGA GTC CGG ATC GCT ATA GGA TCA GCC AAA GGG TTG GCA TAT CTT CAC GAA GAC TGT   1440
 481    H   P   K   I   I   H   R   D   I   K   S   A   N   I   L   L   D   D   E   Y    500
1441  CAT CCG AAA ATC ATT CAC AGA GAT ATA AAG TCA GCA AAT ATT CTT CTA GAT GAT GAA TAT   1500
 501    E   A   Q   A   I   M   K   S   S   F   S   L   N   L   S   Y   D   C   K   V    520
1501  GAA GCT CAG GCA ATA ATG AAA TCC TCC TTT TCG TTA AAT CTA TCT TAT GAC TGT AAA GTT   1560
 521    L   V   A   D   F   G   L   A   R   L   N   D   T   T   Q   T   H   V   S   T    540
1561  TTA GTT GCT GAT TTT GGA CTT GCT AGA CTC AAT GAT ACA ACA CAA ACT CAT GTT TCA ACT   1620
 541    R   V   M   G   T   F   G   Y   L   A   P   E   Y   A   S   S   G   K   L   T    560
1621  CGG GTT ATG GGA ACC TTC GGG TAC CTA GCG CCG GAA TAT GCA TCA AGT GGA AAA TTG ACT   1680
 561    D   R   S   D   V   F   S   F   G   V   V   L   L   E   L   V   T   G   R   K    580
1681  GAT AGA TCC GAT GTA TTC TCA TTC GGG GTT GTT CTC TTA GAG CTT GTA ACT GGA CGG AAA   1740
 581    P   V   D   Q   T   Q   P   L   G   E   E   S   L   V   E   W   A   R   P   L    600
1741  CCA GTT GAC CAG ACT CAG CCT CTA GGA GAA GAG AGT TTG GTT GAA TGG GCG CGC CCG CTG   1800
 601    L   L   K   A   I   E   T   G   D   L   S   E   L   I   D   T   R   L   E   K    620
1801  CTT CTC AAA GCC ATT GAG ACC GGA GAT TTA AGC GAA CTG ATT GAT ACA CGG CTT GAA AAG   1860
 621    R   Y   V   E   H   E   V   F   R   M   I   E   T   A   A   A   C   V   R   H    640
1861  CGT TAT GTG GAG CAT GAA GTC TTC AGA ATG ATC GAG ACA GCC GCT GCA TGT GTT AGA CAT   1920
 641    S   G   P   K   R   P   R   M   V   Q   V   V   R   A   L   D   C   D   G   D    660
1921  TCT GGT CCA AAA CGC CCA CGC ATG GTT CAG GTT GTG AGA GCA TTG GAC TGC GAC GGA GAC   1980
 661    S   G   D   I   S   N   G   I   K   I   G   Q   S   T   T   Y   D   S   G   Q    680
1981  TCG GGA GAT ATT AGC AAC GGA ATC AAA ATT GGG CAA AGC ACA ACT TAT GAC TCA GGG CAA   2040
 681    Y   N   E   D   I   M   K   F   R   K   M   A   F   G   G   D   N   S   V   E    700
2041  TAC AAT GAA GAC ATT ATG AAA TTC AGG AAA ATG GCG TTT GGT GGT GAT AAC AGC GTA GAG   2100
 701    S   G   L   Y   S   G   N   Y   S   A   K   S   S   S   D   F   S   G   N   E    720
2101  TCA GGA TTG TAC AGT GGA AAC TAC TCT GCC AAA AGC TCT TCA GAT TTC TCA GGG AAT GAA   2160
 721    S   E   T   R   P   F   N   N   R   R   F   *                                    732
2161  TCT GAG ACT CGG CCT TTC AAC AAC CGA CGG TTC TGA                                   2196
```

Figure 11B

GAAAATTTTG ATCTCCGATG GCTTCTTCTC CTGAATCTGC TCCTCCAACA AACTCCACCT
CTTCTCCATC TCCACCGTCT AATACCAATT CAACCACCTC TTCTCCGCCG GCTCCGTCTC
CTCCTTCTCC TACACCTCCT CAAGGAGACT CATCATCATC GCCACCTCCT GATTCCACAT
CTCCACCAGC TCCACAAGCT CCTAACCCTC CTAATTCCTC TAATAACTCT CCTTCCCCTC
CGTCACAGGG CGGTGGAGGA GAAAGAGGAA ATGGAGGAAA CAATGGTGGC AATGATACTC
CACCGTCACG CGGCTCTCCT CCTTCTCCTC CTTCTAGGAG TAATGGAGAT AATGGTGGTA
GCAGATCATC GCCACCAGGA GACACTGGAG GCTCTCGCTC AGACAACCCT CCTTCTAGCG
GAGGAAGCAG TGGAGGAGGT GGAGGTGGAA GAAGTAATAC GAATACAGCG ATCATAGTTG
GTGTATTAGT CGGAGCTGGA CTTTTGATGA TCGTTCTTAT TATTGTGTGT CTTAGACGCA
AAAAGAAGAG AAAAGACTCC TTCTACCCTG AACCCATGAA AGGTAAAAAC ATATACACAC
TCTTATGTTT CAACAAATAA GAAGCTTAGA TTCTTTCATA AAATTTCAGG AAACCAATAT
CAATACTATG GAAACAACAA CAACAACAAT GCTTCACAGA ATTATCCGAA TTGGCACCTA
AATTCACAAG GCCAAAACCA ACAATCTACT GGTGGTTGGG GAGGCGGTGG ACCATCACCG
CCTCCTCCTC CGCGGATGCC TACAAGCGGA GAAGATTCTT CCATGTACTC AGGCCCATCA
CGCCCAGTTT TACCTCCTCC TTCGCCTGCT CTAGCCCTCG GATTCAACAA GAGCACTTTT
ACTTACCAAG AGCTTGCGGC TGCAACAGGA GGGTTTACGG ATGCTAACCT TTTGGGACAG
GGAGGATTTG GGTATGTCCA TAAAGGAGTC TTGCCTAGCG GGAAAGAAGT AGCAGTTAAG
AGTTTAAAAG CGGGTAGCGG ACAAGGAGAG AGGGAGTTTC AAGCTGAGGT CGATATCATT
AGCCGTGTGC ATCATCGGTA TCTTGTTTCT TTGGTTGGAT ATTGCATAGC TGATGGACAG
AGGATGTTGG TTTATGAGTT TGTTCCTAAC AAAACTTTGG AATATCATCT TCATGGTTAG
ACCACTTAAA AACTTTGAGT ACTAAGTTTA TTTTCTCTAA TCTATATATT CAAGAAAGTT
GTAACCTTAA TTTGTTGTTG TAGGGAAAAA TCTTCCGGTA ATGGAGTTCT CCACTAGGTT
GCGTATCGCC TTAGGTGCTG CGAAAGGACT CGCTTACCTT CACGAAGACT GTAAGTTTTA
ACATTCACCA TTCTCATTTT CTTAACCAAG TTGCATAAAA CAGAGAAAGC TCTGTCTCTG
ACTAGTGTTA TCTTTTTGGC TGAGAAAATG GTGCAGGCCA TCCTCGGATC ATTCACCGCG
ACATCAAGTC TGCAAATATT CTCTTGGACT TCAACTTTGA TGCTATGGTG ATAAACTAGT
AGCTTGCATT CATCTACGGT TTTTTGTTAA GACTACATTG ATGACATTTT GCATTTGTTT
ATTCAGGTGG CTGATTTTGG ATTAGCTAAG TTAACATCTG ATAACAACAC TCATGTATCT
ACTCGTGTGA TGGGAACTTT CGGGTAAGCG TTTTACCGTA TGATAAGATT GTTCGTGACA
CTCAAGAAAC ATAACCTTTG TAGACTAATC TACTTTGCTT CTTTCCACAA ACATGTGTAG
ATATCTAGCT CCAGAATATG CTTCAAGCGG TAAATTAACC GAGAAATCCG ATGTTTTCTC
TTACGGAGTT ATGTTATTGG AACTTATAAC TGGAAAACGA CCGGTTGATA ATAGCATCAC
CATGGACGAC ACCTTAGTAG ATTGGGTATT CATGCATGTA ACATATGTAT CGTGTATATA
TGTTTTTCGC CTTTTTCGCG TACTAATGAT CATGAATACA GGCTCGGCCT CTTATGGCTC
GCGCGCTAGA AGATGGAAAC TTTAATGAGC TCGCAGATGC GAGGCTTGAA GGCAACTACA
ACCCGCAAGA AATGGCTCGA ATGGTGACTT GTGCCGCTGC TAGCATTCGT CATTCGGGGC
GTAAACGTCC AAAGATGAGC CAGGTGAATC AAAATTATAA CTAAAGTCT ATTTTTGTCA
GAGAATAACA AACAAATGTT GTGGTTTTCA GATAGTAAGA GCGTTAGAAG GAGAAGTGTC
CTTAGATGCT TTAAACGAAG GTGTGAAGCC AGGACACAGT AACGTTTACG GGTCATTGGG
AGCAAGCTCG GATTATAGTC AGACATCTTA CAATGCAGAC ATGAAGAAAT TCAGACAGAT
AGCTTTGTCG AGCCAAGAAT TCCCAGTCAG TGACTGTGAA GGAACATCTA GTAATGATTC
TAGAGATATG GGAACTAAGA GCCCTACTCC TCCAAAATGA GATCGAATCA ATGATTCTGT

Figure 12A

```
   1   M   A   S   S   P   E   S   A   P   P   T   N   S   T   S   S   P   S   P   P    20
   1  ATG GCT TCT TCT CCT GAA TCT GCT CCT CCA ACA AAC TCC ACC TCT TCT CCA TCT CCA CCG   60

  21   S   N   T   N   S   T   T   S   S   P   P   A   P   S   P   P   S   P   T   P    40
  61  TCT AAT ACC AAT TCA ACC ACC TCT TCT CCG CCG GCT CCG TCT CCT CCT TCT CCT ACA CCT  120

  41   P   Q   G   D   S   S   S   S   P   P   P   D   S   T   S   P   P   A   P   Q    60
 121  CCT CAA GGA GAC TCA TCA TCA TCG CCA CCT CCT GAT TCC ACA TCT CCA CCA GCT CCA CAA  180

  61   A   P   N   P   P   N   S   S   N   N   S   P   S   P   P   S   Q   G   G   G    80
 181  GCT CCT AAC CCT CCT AAT TCC TCT AAT AAC TCT CCT TCC CCT CCG TCA CAG GGC GGT GGA  240

  81   G   E   R   G   N   G   G   N   N   G   G   N   D   T   P   P   S   R   G   S   100
 241  GGA GAA AGA GGA AAT GGA GGA AAC AAT GGT GGC AAT GAT ACT CCA CCG TCA CGC GGC TCT  300

 101   P   P   S   P   P   S   R   S   N   G   D   N   G   G   S   R   S   S   P   P   120
 301  CCT CCT TCT CCT CCT TCT AGG AGT AAT GGA GAT AAT GGT GGT AGC AGA TCA TCG CCA CCA  360

 121   G   D   T   G   G   S   R   S   D   N   P   P   S   S   G   G   S   S   G   G   140
 361  GGA GAC ACT GGA GGC TCT CGC TCA GAC AAC CCT CCT TCT AGC GGA GGA AGC AGT GGA GGA  420

 141   G   G   G   G   R   S   N   T   N   T   A   I   I   V   G   V   L   V   G   A   160
                                           ‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾
 421  GGT GGA GGT GGA AGA AGT AAT ACG AAT ACA GCG ATC ATA GTT GGT GTA TTA GTC GGA GCT  480

 161   G   L   L   M   I   V   L   I   I   V   C   L   R   R   K   K   K   R   K   D   180
       ‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾
 481  GGA CTT TTG ATG ATC GTT CTT ATT ATT GTG TGT CTT AGA CGC AAA AAG AAG AGA AAA GAC  540

 181   S   F   Y   P   E   P   M   K   G   N   Q   Y   Q   Y   Y   G   N   N   N   N   200
 541  TCC TTC TAC CCT GAA CCC ATG AAA GGA AAC CAA TAT CAA TAC TAT GGA AAC AAC AAC AAC  600

 201   N   N   A   S   Q   N   Y   P   N   W   H   L   N   S   Q   G   Q   N   Q   Q   220
 601  AAC AAT GCT TCA CAG AAT TAT CCG AAT TGG CAC CTA AAT TCA CAA GGC CAA AAC CAA CAA  660

 221   S   T   G   G   W   G   G   G   G   P   S   P   P   P   P   P   R   M   P   T   240
 661  TCT ACT GGT GGT TGG GGA GGC GGT GGA CCA TCA CCG CCT CCT CCT CCG CGG ATG CCT ACA  720

 241   S   G   E   D   S   S   M   Y   S   G   P   S   R   P   V   L   P   P   P   S   260
 721  AGC GGA GAA GAT TCT TCC ATG TAC TCA GGC CCA TCA CGC CCA GTT TTA CCT CCT CCT TCG  780

 261   P   A   L   A   L   G   F   N   K   S   T   F   T   Y   Q   E   L   A   A   A   280
 781  CCT GCT CTA GCC CTC GGA TTC AAC AAG AGC ACT TTT ACT TAC CAA GAG CTT GCG GCT GCA  840

 281   T   G   G   F   T   D   A   N   L   L   G   Q   G   G   F   G   Y   V   H   K   300
 841  ACA GGA GGG TTT ACG GAT GCT AAC CTT TTG GGA CAG GGA GGA TTT GGG TAT GTC CAT AAA  900

 301   G   V   L   P   S   G   K   E   V   A   V   K   S   L   K   A   G   S   G   Q   320
 901  GGA GTC TTG CCT AGC GGG AAA GAA GTA GCA GTT AAG AGT TTA AAA GCG GGT AGC GGA CAA  960

 321   G   E   R   E   F   Q   A   E   V   D   I   I   S   R   V   H   H   R   Y   L   340
 961  GGA GAG AGG GAG TTT CAA GCT GAG GTC GAT ATC ATT AGC CGT GTG CAT CAT CGG TAT CTT 1020

 341   V   S   L   V   G   Y   C   I   A   D   G   Q   R   M   L   V   Y   E   F   V   360
1021  GTT TCT TTG GTT GGA TAT TGC ATA GCT GAT GGA CAG AGG ATG TTG GTT TAT GAG TTT GTT 1080

 361   P   N   K   T   L   E   Y   H   L   H   G   K   N   L   P   V   M   E   F   S   380
1081  CCT AAC AAA ACT TTG GAA TAT CAT CTT CAT GGG AAA AAT CTT CCG GTA ATG GAG TTC TCC 1140

 381   T   R   L   R   I   A   L   G   A   A   K   G   L   A   Y   L   H   E   D   C   400
1141  ACT AGG TTG CGT ATC GCC TTA GGT GCT GCG AAA GGA CTC GCT TAC CTT CAC GAA GAC TGC 1200

 401   H   P   R   I   I   H   R   D   I   K   S   A   N   I   L   L   D   F   N   F   420
1201  CAT CCT CGG ATC ATT CAC CGC GAC ATC AAG TCT GCA AAT ATT CTC TTG GAC TTC AAC TTT 1260

 421   D   A   M   V   A   D   F   G   L   A   K   L   T   S   D   N   N   T   H   V   440
1261  GAT GCT ATG GTG GCT GAT TTT GGA TTA GCT AAG TTA ACA TCT GAT AAC AAC ACT CAT GTA 1320

 441   S   T   R   V   M   G   T   F   G   Y   L   A   P   E   Y   A   S   S   G   K   460
1321  TCT ACT CGT GTG ATG GGA ACT TTC GGA TAT CTA GCT CCA GAA TAT GCT TCA AGC GGT AAA 1380

 461   L   T   E   K   S   D   V   F   S   Y   G   V   M   L   L   E   L   I   T   G   480
1381  TTA ACC GAG AAA TCC GAT GTT TTC TCT TAC GGA GTT ATG TTA TTG GAA CTT ATA ACT GGA 1440

 481   K   R   P   V   D   N   S   I   T   M   D   D   T   L   V   D   W   A   R   P   500
1441  AAA CGA CCG GTT GAT AAT AGC ATC ACC ATG GAC GAC ACC TTA GTA GAT TGG GCT CGG CCT 1500

 501   L   M   A   R   A   L   E   D   G   N   F   N   E   L   A   D   A   R   L   E   520
1501  CTT ATG GCT CGC GCG CTA GAA GAT GGA AAC TTT AAT GAG CTC GCA GAT GCG AGG CTT GAA 1560

 521   G   N   Y   N   P   Q   E   M   A   R   M   V   T   C   A   A   A   S   I   R   540
1561  GGC AAC TAC AAC CCG CAA GAA ATG GCT CGA ATG GTG ACT TGT GCC GCT GCT AGC ATT CGT 1620

 541   H   S   G   R   K   R   P   K   M   S   Q   I   V   R   A   L   E   G   E   V   560
1621  CAT TCG GGG CGT AAA CGT CCA AAG ATG AGC CAG ATA GTA AGA GCG TTA GAA GGA GAA GTG 1680

 561   S   L   D   A   L   N   E   G   V   K   P   G   H   S   N   V   Y   G   S   L   580
1681  TCC TTA GAT GCT TTA AAC GAA GGT GTG AAG CCA GGA CAC AGT AAC GTT TAC GGG TCA TTG 1740

 581   G   A   S   S   D   Y   S   Q   T   S   Y   N   A   D   M   K   K   F   R   Q   600
1741  GGA GCA AGC TCG GAT TAT AGT CAG ACA TCT TAC AAT GCA GAC ATG AAG AAA TTC AGA CAG 1800

 601   I   A   L   S   S   Q   E   F   P   V   S   D   C   E   G   T   S   S   N   D   620
1801  ATA GCT TTG TCG AGC CAA GAA TTC CCA GTC AGT GAC TGT GAA GGA ACA TCT AGT AAT GAT 1860

 621   S   R   D   M   G   T   K   S   P   T   P   P   K   *                          634
1861  TCT AGA GAT ATG GGA ACT AAG AGC CCT ACT CCT CCA AAA TGA                         1902
```

Figure 12B

TCCACCGTTT GAGAAACCCT AATAACAACA TTCAAAATGG CGGACTCACC GGTGGATTCA

TCTCCTGCCC CTGAAACCTC AAATGGGACA CCACCGTCAA ACGGAACATC GCCGTCTAAT

GAGTCATCGC CGCCAACACC ACCTTCTTCA CCACCACCAT CATCAATATC TGCTCCTCCG

CCAGATATCT CCGCTTCTTT TTCACCGCCG CCTGCACCAC CAACGCAAGA AACGTCACCT

CCTACATCTC CGTCCTCATC GCCGCCTGTT GTAGCTAATC CGTCACCGCA GACTCCAGAG

AATCCTTCTC CACCTGCACC TGAAGGCTCA ACTCCTGTAA CGCCACCTGC ACCACCACAA

ACACCGTCGA ACCAATCACC GGAAAGACCA ACTCCTCCTT CTCCTGGTGC CAATGATGAC

CGAAACAGAA CCAATGGCGG AAACAACAAC AGAGACGGCT CCACACCATC ACCACCGTCG

TCAGGGAACA GAACTTCCGG TGACGGTGGC TCACCTTCAC CACCTCGGTC GATAAGCCCT

CCTCAGAATA GTGGAGATTC AGACTCATCA TCGGGTAATC ATCCACAAGC CAACATTGGA

TTGATTATTG GAGTCCTTGT AGGAGCAGGG CTTTTGCTTC TACTTGCAGT GTGTATTTGC

ATCTGTTGCA ACAGGAAGAA GAAGAAGAAA TCTCCTCAGG TCAACCACAT GCACTACTAC

AATAACAATC CTTATGGAGG AGCACCCTCA GGTAATTACA GTTTAGTATA ACTGGAATTT

AATTTGTAGC CTAATGGTGT TGATTAGGT TTCAGAACGA TCATAGTCTA ATGGTTTCTG

CTAGCTCCAT ATGGCAAAAG GATTAGATTT ATAAGCTAAA GGAGATGTTG CATAGTGTAG

GTAATGGTGG TTATTACAAG GGAACACCTC AAGATCATGT GGTGAATATG GCTGGTCAAG

GAGGTGGGAA TTGGGGTCCA CAGCAACCTG TGTCTGGTCC TCACAGTGAT GCTTCCAACT

TAACCGGTCG AACTGCTATA CCGTCACCTC AAGCTGCAAC TCTTGGTCAC AACCAAAGCA

CTTTCACATA CGATGAACTG TCCATTGCAA CAGAAGGTTT CGCTCAGTCA AATTTGCTAG

GACAAGGAGG ATTTGGGTAT GTTCATAAAG GAGTTCTGCC TAGTGGCAAA GAAGTTGCAG

TGAAGAGTCT TAAACTTGGA AGTGGACAAG GGGAACGCGA GTTTCAAGCA GAGGTTGATA

TCATTAGCCG TGTCCATCAT CGTCATCTCG TTTCTCTTGT TGGATATTGC ATCTCTGGTG

GTCAAAGACT TTTGGTTTAT GAGTTTATAC CTAACAACAC TCTTGAATTT CATCTTCATG

GTACATTCAT CTAACAGAAT GTTTTCTTGT ATTAACAAAA CCTTTAAGTA TGGTTTCTCT

TTAATCAGGA ACATGATTGA AATTTCAGGA AAGGGTCGTC CGGTTTTGGA TTGGCCTACA

AGAGTGAAGA TTGCATTGGG ATCAGCTAGA GGCCTTGCAT ATTTGCATGA AGACTGTAAG

AAAATCTTTA TCTCACATAT TTGCATCAGT TTCTATCTCG CTCTCTACAA TATTTGAAAG

ATTGTATATG TTACATCAAT TATAGGTCAC CCTCGCATTA TCCACAGAGA TATCAAAGCT

GCAAACATTC TTCTTGATTT CAGTTTTGAG ACCAAGGTAT GTGTGTATAT ATCGACTCTT

GTACTACTTT TACTTTCATT GTCTCTCATT TTTGTTTCCA ATCTGTGTCG ATGTGTGTAT

CAGTCTTATT GTGTAAATAT ATGCAGGTGG CAGATTTTGG ATTGGCTAAG CTATCTCAAG

ACAACTATAC TCATGTCTCC ACTCGCGTCA TGGGAACTTT TGGGTAAGCA GCTTTGTAAA

ATGTCTCAAC TCATCCACAC TTATTTAGTT TCTTTCACTT GTTTTTAACA TTTTCTTGGA

TTCAGATACT TAGCTCCAGA GTATGCATCA AGCGGAAAGT TATCCGACAA ATCTGATGTT

TTCTCATTTG GAGTAATGCT TCTTGAGCTC ATAACCGGAA GACCTCCTCT GGATCTAACT

GGAGAAATGG AAGATAGCTT GGTAGATTGG GTAAGTCGGT CCCCGCCTCT TCGGTTTACT

TGTTTAATCC CAAAACACTT TCCAAAGCAA AAACAGAAAC AAATCTTACT ATTGTTGTTG

CAGGCAAGGC CTTTGTGTTT GAAAGCAGCT CAAGATGGAG ATTACAACCA ATTGGCTGAT

CCACGTCTAG AGCTAAACTA CAGTCATCAA GAGATGGTTC AAATGGCTTC TTGTGCAGCT

GCAGCAATCA GACATTCAGC AAGAAGACGG CCTAAGATGA GCCAGGTTCA AAAACTCATA

CCACTTGTTG GTTCTATTGT TATATTTTTA CTCACAATTA ATCTTGATGA TAAATGTGAC

ATACTAATGA ATCTTGAAAC ATGTGTATGG TAAATGAAAA GATTGTACGA GCACTAGAAG

GAGATATGTC AATGGATGAT CTAAGTGAGG GAACAAGACC AGGACAAAGC ACGTACTTGA

GCCCCGGGAG CGTGAGCTCA GAGTATGACG CAAGCTCGTA CACGGCAGAC ATGAAAAAGT

TCAAGAAACT GGCGTTAGAG AATAAAGAAT ATCAAAGCAG TGAATATGGT GGAACAAGTG

AGTATGGCTT AAACCCTTCT GCTTCAAGTA GTGAAGAAAT GAATAGAGGC TCAATGAAAC

GCAATCCTCA GCTTTGAAAG AAGAGACAAC ACTTGTCATA ATATTTCAGT TTTCTTCTCT

Figure 13A

```
  1    M   A   D   S   P   V   D   S   S   P   A   P   E   T   S   N   G   T   P   P    20
  1   ATG GCG GAC TCA CCG GTG GAT TCA TCT CCT GCC CCT GAA ACC TCA AAT GGG ACA CCA CCG   60
 21    S   N   G   T   S   P   S   N   E   S   S   P   P   T   P   P   S   S   S   P    40
 61   TCA AAC GGA ACA TCG CCG TCT AAT GAG TCA TCG CCG CCA ACA CCA CCT TCT TCA CCA CCA  120
 41    P   S   S   I   S   A   P   P   P   D   I   S   A   S   F   S   P   P   P   A    60
121   CCA TCA TCA ATA TCT GCT CCT CCG CCA GAT ATC TCC GCT TCT TTT TCA CCG CCG CCT GCA  180
 61    P   P   T   Q   E   T   S   P   P   T   S   P   S   S   S   P   P   V   V   A    80
181   CCA CCA ACG CAA GAA ACG TCA CCT CCT ACA TCT CCG TCC TCA TCG CCG CCT GTT GTA GCT  240
 81    N   P   S   P   Q   T   P   E   N   P   S   P   A   P   E   G   S   T   P       100
241   AAT CCG TCA CCG CAG ACT CCA GAG AAT CCT TCT CCA CCT GCA CCT GAA GGC TCA ACT CCT  300
101    V   T   P   P   A   P   P   Q   T   P   S   N   Q   S   P   E   R   P   T   P   120
301   GTA ACG CCA CCT GCA CCA CCA CAA ACA CCG TCG AAC CAA TCA CCG GAA AGA CCA ACT CCT  360
121    P   S   P   G   A   N   D   D   R   N   R   T   N   G   G   N   N   N   R   D   140
361   CCT TCT CCT GGT GCC AAT GAT GAC CGA AAC AGA ACC AAT GGC GGA AAC AAC AAC AGA GAC  420
141    G   S   T   P   S   P   P   S   S   G   N   R   T   S   G   D   G   G   S   P   160
421   GGC TCC ACA CCA TCA CCA CCG TCG TCA GGG AAC AGA ACT TCC GGT GAC GGT GGC TCA CCT  480
161    S   P   P   R   S   I   S   P   P   Q   N   S   G   D   S   D   S   S   S   G_  180
481   TCA CCA CCT CGG TCG ATA AGC CCT CCT CAG AAT AGT GGA GAT TCA GAC TCA TCA TCG GGG  540
181   _L   L   L   L   L   A   V   C   I   C   I   C   C   N_  R   K   K   K   K   K   200
541   CTT TTG CTT CTA CTT GCA GTG TGT ATT TGC ATC TGT TGC AAC AGG AAG AAG AAG AAG AAA  600
201    S   P   Q   V   N   H   M   H   Y   Y   N   N   N   P   Y   G   G   A   P   S   220
601   TCT CCT CAG GTC AAC CAC ATG CAC TAC TAC AAT AAC AAT CCT TAT GGA GGA GCA CCC TCA  660
221    G   N   G   G   Y   Y   K   G   T   P   Q   D   H   V   V   N   M   A   G   Q   240
661   GGT AAT GGT GGT TAT TAC AAG GGA ACA CCT CAA GAT CAT GTG GTG AAT ATG GCT GGT CAA  720
241    G   G   G   N   W   G   P   Q   Q   P   V   S   G   P   H   S   D   A   S   N   260
721   GGA GGT GGG AAT TGG GGT CCA CAG CAA CCT GTG TCT GGT CCT CAC AGT GAT GCT TCC AAC  780
261    L   T   G   R   T   A   I   P   S   P   Q   A   A   T   L   G   H   N   Q   S   280
781   TTA ACC GGT CGA ACT GCT ATA CCG TCA CCT CAA GCT GCA ACT CTT GGT CAC AAC CAA AGC  840
281    T   F   T   Y   D   E   L   S   I   A   T   E   G   F   A   Q   S   N   L   L   300
841   ACT TTC ACA TAC GAT GAA CTG TCC ATT GCA ACA GAA GGT TTC GCT CAG TCA AAT TTG CTA  900
301    G   Q   G   G   F   G   Y   V   H   K   G   V   L   P   S   G   K   E   V   A   320
901   GGA CAA GGA GGA TTT GGG TAT GTT CAT AAA GGA GTT CTG CCT TCA GGC AAA GAA GTT GCA  960
321    V   K   S   L   K   L   G   S   G   Q   G   E   R   E   F   Q   A   E   V   D   340
961   GTG AAG AGT CTT AAA CTT GGA AGT GGA CAA GGG GAA CGC GAG TTT CAA GCA GAG GTT GAT 1020
341    I   I   S   R   V   H   H   R   H   L   V   S   L   V   G   Y   C   I   S   G   360
1021  ATC ATT AGC CGT GTC CAT CAT CGT CAT CTC GTT TCT CTT GTT GGA TAT TGC ATC TCT GGT 1080
361    G   Q   R   L   L   V   Y   E   F   I   P   N   N   T   L   E   F   H   L   H   380
1081  GGT CAA AGA CTT TTG GTT TAT GAG TTT ATA CCT AAC AAC ACT CTT GAA TTT CAT CTT CAT 1140
381    G   K   G   R   P   V   L   D   W   P   T   R   V   K   I   A   L   G   S   A   400
1141  GGA AAG GGT CGT CCG GTT TTG GAT TGG CCT ACA AGA GTG AAG ATT GCA TTG GGA TCA GCT 1200
401    R   G   L   A   Y   L   H   E   D   C   K   K   I   F   I   S   H   I   C   I   420
1201  AGA GGC CTT GCA TAT TTG CAT GAA GAC TGT AAG AAA ATC TTT ATC TCA CAT ATT TGC ATC 1260
421    S   H   P   R   I   I   H   R   D   I   K   A   A   N   I   L   L   D   F   S   440
1261  AGT CAC CCT CGC ATT ATC CAC AGA GAT ATC AAA GCT GCA AAC ATT CTT CTT GAT TTC AGT 1320
441    F   E   T   K   V   A   D   F   G   L   A   K   L   S   Q   D   N   Y   T   H   460
1321  TTT GAG ACC AAG GTG GCA GAT TTT GGA TTG GCT AAG CTA TCT CAA GAC AAC TAT ACT CAT 1380
461    V   S   T   R   V   M   G   T   F   G   Y   L   A   P   E   Y   A   S   S   G   480
1381  GTC TCC ACT CGC GTC ATG GGA ACT TTT GGA TAC TTA GCT CCA GAG TAT GCA TCA AGC GGA 1440
481    K   L   S   D   K   S   D   V   F   S   F   G   V   M   L   L   E   L   I   T   500
1441  AAG TTA TCC GAC AAA TCT GAT GTT TTC TCA TTT GGA GTA ATG CTT CTT GAG CTC ATA ACC 1500
501    G   R   P   P   L   D   L   T   G   E   M   E   D   S   L   V   D   W   A   R   520
1501  GGA AGA CCT CCT CTG GAT CTA ACT GGA GAA ATG GAA GAT AGC TTG GTA GAT TGG GCA AGG 1560
521    P   L   C   L   K   A   A   Q   D   G   D   Y   N   Q   L   A   D   P   R   L   540
1561  CCT TTG TGT TTG AAA GCA GCT CAA GAT GGA GAT TAC AAC CAA TTG GCT GAT CCA CGT CTA 1620
541    E   L   N   Y   S   H   Q   E   M   V   Q   M   A   S   C   A   A   A   A   I   560
1621  GAG CTA AAC TAC AGT CAT CAA GAG ATG GTT CAA ATG GCT TCT TGT GCA GCT GCA GCA ATC 1680
561    R   H   S   A   R   R   R   P   K   M   S   Q   V   Q   K   L   I   P   L   V   580
1681  AGA CAT TCA GCA AGA AGA CGG CCT AAG ATG AGC CAG GTT CAA AAA CTC ATA CCA CTT GTT 1740
581    G   S   I   I   V   R   A   L   E   G   D   M   S   M   D   D   L   S   E   G   600
1741  GGT TCT ATT ATT GTA CGA GCA CTA GAA GGA GAT ATG TCA ATG GAT GAT CTA AGT GAG GGA 1800
601    T   R   P   G   Q   S   T   Y   L   S   P   G   S   V   S   S   E   Y   D   A   620
1801  ACA AGA CCA GGA CAA AGC ACG TAC TTG AGC CCC GGG AGC GTG AGC TCA GAG TAT GAC GCA 1860
621    S   S   Y   T   A   D   M   K   K   F   K   K   L   A   L   E   N   K   E   Y   640
1861  AGC TCG TAC ACG GCA GAC ATG AAA AAG TTC AAG AAA CTG GCG TTA GAG AAT AAA GAA TAT 1920
641    Q   S   S   E   Y   G   G   T   S   E   Y   G   L   N   P   S   A   S   S       660
1921  CAA AGC AGT GAA TAT GGT GGA ACA AGT GAG TAT GGC TTA AAC CCT TCT GCT TCA AGT AGT 1980
661    E   E   M   N   R   G   S   M   K   R   N   P   Q   L   *                       675
1981  GAA GAA ATG AAT AGA GGC TCA ATG AAA CGC AAT CCT CAG CTT TGA                      2025
```

<div align="center">Figure 13B</div>

CATACATCAC AAAACGGCAT CGTTTTGATG TCGCTCTCTC CTTCTTCCTC GCCGGCTCCG
GCAACTTCCC CGCCTGCCAT GTCATTACCT CCGGCAGATT CCGTACCTGA CACGTCATCA
CCTCCAGCTC CTCCTTTGTC TCCTCTTCCC CCACCATTGA GCTCTCCTCC GCCGTTGCCT
TCACCACCGC CTCTCTCCGC TCCCACCGCT TCCCCACCGC CTCTTCCGGT TGAATCCCCA
CCGTCTCCTC CTATAGAATC ACCACCGCCT CCTCTACTGG AATCACCTCC TCCTCCTCCG
TTGGAATCTC CATCTCCACC GTCTCCTCAC GTCTCAGCTC CTTCCGGTTC ACCGCCATTA
CCCTTCCTTC CCGCCAAACC TTCTCCGCCG CCTTCTTCAC CTCCCTCCGA GACAGTTCCG
CCGGGAAATA CGATTTCTCC ACCACCTCGT TCACTTCCCT CCGAATCAAC CCCGCCGGTG
AACACAGCTT CTCCTCCACC GCCATCTCCT CCTCGCCGCC GTAGTGGCCC TAAGCCTTCG
TTTCCTCCTC CCATCAATTC TTCTCCACCA AATCCTTCTC CGAACACTCC GTCACTCCCA
GAAACTTCTC CTCCACCTAA ACCACCGCTC TCAACGACGC CATTTCCCTC CTCATCCACT
CCCCCGCCTA AGAAGTCCCC TGCAGCAGTA ACTCTTCCTT TCTTTGGGCC AGCGGGCCAA
TTACCGGATG GGACCGTAGC ACCTCCTATT GGGCCTGTTA TAGAGCCCAA GACGAGTCCA
GCCGAATCAA TATCTCCGGG AACGCCACAG CCACTGGTTC CGAAGAGTCT ACCTGTAACG
ACGTCGTATC ACCGATCATC CGCCGGATTC TTATTTGGCG GTGTAATCGT TGGAGCTCTT
CTACTAATTC TGTTAGGTCT TCTCTTTGTC TTCTACAGAG CTACCAGAAA TAGAAATAAC
AACAGCAGCT CTGCTCATCA TCAATCCAAA ACTCCCTCAA AAGGTATAAA CTTTGAGATC
AATTTGTTTC AGACTGTCAC TTAAAATGTA ATGTAATTGT AAGTTATGAT CTTTTTTGCT
TGGATCAGTT CAACATCATC GGGGCGGTAA TGCTGGTACG AACCAGGCAC ATGTTATCAC
AATGCCACCA CCAATCCATG CTAAATATAT ATCTAGTGGA GGTTGTGATA CGAAGGAGAA
CAATTCTGTT GCGAAAAACA TTTCAATGCC ATCTGGAATG TTCTCCTACG AAGAACTTTC
AAAAGCAACT GGTGGATTTT CAGAGGAGAA CCTTTTGGGA GAAGGCGGTT TCGGATATGT
TCACAAAGGA GTGTTGAAAA ACGGGACAGA AGTTGCGGTG AAGCAGCTGA AGATTGGGAG
CTATCAAGGG GAAAGAGAAT TCCAAGCTGA GGTTGACACA ATCAGTAGGG TTCATCATAA
GCACCTCGTT TCATTGGTTG GTTATTGCGT TAATGGAGAT AAAAGACTCT TGGTTTACGA
GTTTGTTCCT AAAGATACCT TGGAGTTCCA CTTGCATGGT AAAATAGATA TATGATTTCA
TCCTTTTTGA TTTTGTCTTT AGTTCATTAT GTTGAGTATT GTGAGAATAT GTGTTGTAGA
GAACAGAGGA AGCGTGTTGG AATGGGAAAT GAGGCTCAGG ATTGCTGTAG GAGCAGCAAA
AGGATTAGCT TATCTTCATG AGGATTGTGA GTTGTTTCTC TTCATAATGG GAATGACAAT
GGCCTTTTTT GGGTTTTCGA CACTGATATT GATTTCTGGT GCTTGTAATG CAGGCAGTCC
AACTATAATT CACCGTGATA TCAAAGCAGC TAATATCCTT CTAGATTCCA AATTTGAGGC
AAAGGTGATC TGCTCTTTAA TCTTATCAAA GTTTGGTTTT TAGAACAGAG TTTGCCCATT
TTCGGTTTAT AGTCACACCA TTTTGTTCTA TTTAACAAGA TTACATGTAG CTTAGAGTTC
TTGTTCTCTT CAGGTCTCTG ACTTTGGACT AGCCAAGTTT TTCTCAGACA CCAATTCATC
ATTCACTCAT ATCTCTACTC GAGTGGTAGG AACTTTCGGG TAAAACACCA TCCATCCATG
CTTTATATGT TGTGTCATTG TGTTTAAAAA TATTAATTTA CGGTTTCAAG TTTCAAGTTT
CAAGTTTCAA GTTTCTTCCT TGTTGTATCG TAAATCCAGA TCAAAGATT TATTGATTAC
TAAATGCCTT GTGCACCGTC TATTTGGTAT GCATTTAAAC AGATACATGG CTCCAGAATA
CGCGTCCAGT GGTAAAGTAA CTGATAAATC AGATGTATAT TCCTTTGGGG TCGTGCTTCT
AGAACTCATC ACTGGACGTC CATCAATTTT CGCCAAAGAT TCTTCCACAA ACCAGAGTTT
AGTAGACTGG GTAAGTCAAA GTACATGATG ATGATGATGA TACCATTAGG TTTTCTTTTT
CCCACTAGTT ATTAGATGAA AATGAAACTC CAGTATAACT GCAAGCTTGA AACTTCATTA
GACTGTAAAT TTTGATTATT CCTCAGGCGA GGCCATTGCT TACGAAAGCA ATCTCTGGAG
AAAGTTTTGA CTTTCTTGTA GACTCAAGGT TGGAGAAGAA TTACGATACA ACTCAGATGG
CAAACATGGC TGCTTGTGCT GCTGCTTGCA TACGCCAATC AGCTTGGCTT CGGCCTAGAA
TGAGCCAGGT CTGAGATTTA GTTAAATAC ATGTATTCGT CCATATATCC GAAAGGAACA
AGCCTAATCC ATGAATACAT TTATATCTTG AAAAAAACTT GAATAGGTAG TACGTGCTCT
TGAAGGCGAG GTGGCCCTGA GAAAGGTCGA AGAGACTGGG AATAGCGTGA CCTATAGCTC
TTCTGAAAAC CCGAATGACA TCACACCACG GTATGGAACA AATAAGAGGA GATTCGACAC
AGGTTCAAGC GATGGTTACA CTTCAGAATA TGGAGTTAAC CCTTCTCAGT CGAGCAGTGA
ACATCAACAG GTGAATACTT AGTTCACAGG TTCAATAGGG CAAGTTTCAC CACAATTATT

Figure 14A

```
  1    M   S   L   S   P   S   S   S   P   A   P   A   T   S   P   P   A   M   S   L    20
  1   ATG TCG CTC TCT CCT TCT TCC TCG CCG GCT CCG GCA ACT TCC CCG CCT GCC ATG TCA TTA   60
 21    P   P   A   D   S   V   P   D   T   S   S   P   P   A   P   P   L   S   P   L    40
 61   CCT CCG GCA GAT TCC GTA CCT GAC ACG TCA TCA CCT CCA GCT CCT CCT TTG TCT CCT CTT  120
 61    P   P   P   L   S   S   P   P   P   L   P   S   P   P   P   L   S   A   P   T    80
121   CCC CCA CCA TTG AGC TCT CCT CCG CCG TTG CCT TCA CCA CCG CCT CTC TCC GCT CCC ACC  180
 61    A   S   P   P   P   L   P   V   E   S   P   P   S   P   P   I   E   S   P   P    80
181   GCT TCC CCA CCG CCT CTT CCG GTT GAA TCC CCA CCG TCT CCT CCT ATA GAA TCA CCA CCG  240
 81    P   P   L   L   E   S   P   P   P   P   L   E   S   P   S   P   P   S   P   100
241   CCT CCT CTA CTG GAA TCA CCT CCT CCT CCT CCG TTG GAA TCT CCA TCT CCA CCG TCT CCT  300
101    H   V   S   A   P   S   G   S   P   P   L   P   F   L   P   A   K   P   S   P   120
301   CAC GTC TCA GCT CCT TCC GGT TCA CCG CCA TTA CCC TTC CTT CCC GCC AAA CCT TCT CCG  360
121    P   P   S   S   P   P   S   E   T   V   P   P   G   N   T   I   S   P   P   P   140
361   CCG CCT TCT TCA CCT CCC TCC GAG ACA GTT CCG CCG GGA AAT ACG ATT TCT CCA CCA CCT  420
141    R   S   L   P   S   E   S   T   P   P   V   N   T   A   S   P   P   P   P   S   160
421   CGT TCA CTT CCC TCC GAA TCA ACC CCG CCG GTG AAC ACA GCT TCT CCT CCA CCG CCA TCT  480
161    P   P   R   R   R   S   G   P   K   P   S   F   P   P   P   I   N   S   S   P   180
481   CCT CCT CGC CGC CGT AGT GGC CCT AAG CCT TCG TTT CCT CCT CCC ATC AAT TCT TCT CCA  540
181    P   N   P   S   P   N   T   P   S   L   P   E   T   S   P   P   P   K   P   P   200
541   CCA AAT CCT TCT CCG AAC ACT CCG TCA CTC CCA GAA ACT TCT CCT CCA CCT AAA CCA CCG  600
201    L   S   T   T   P   F   P   S   S   S   T   P   P   P   K   K   S   P   A   A   220
601   CTC TCA ACG ACG CCA TTT CCC TCC TCA TCC ACT CCC CCG CCT AAG AAG TCC CCT GCA GCA  660
221    V   T   L   P   F   F   G   P   A   G   Q   L   P   D   G   T   V   A   P   P   240
661   GTA ACT CTT CCT TTC TTT GGG CCA GCG GGC CAA TTA CCG GAT GGG ACC GTA GCA CCT CCT  720
241    I   G   P   V   I   E   P   K   T   S   P   A   E   S   I   S   P   G   T   P   260
721   ATT GGG CCT GTT ATA GAG CCC AAG ACG AGT CCA GCC GAA TCA ATA TCT CCG GGA ACG CCA  780
261    Q   P   L   V   P   K   S   L   P   V   T   T   S   Y   H   R   S   S   A   G   280
781   CAG CCA CTG GTT CCG AAG AGT CTA CCT GTA ACG TCG TAT CAC CGA TCA TCC GCC GGA  840
281    F   L   F   G   G   V   I   V   G   A   L   L   L   I   L   L   G   L   L   F   300
841   TTC TTA TTT GGC GGT GTA ATC GTT GGA GCT CTT CTA CTA ATT CTG TTA GGT CTT CTC TTT  900
301    V   F   Y   R   A   T   R   N   R   N   N   N   S   S   S   A   H   H   Q   S   320
901   GTC TTC TAC AGA GCT ACC AGA AAT AGA AAT AAC AAC AGC AGC TCT GCT CAT CAT CAA TCC  960
321    K   T   P   S   K   V   Q   H   H   R   G   G   N   A   G   T   N   Q   A   H   340
961   AAA ACT CCC TCA AAA GTT CAA CAT CAT CGG GGC GGT AAT GCT GGT ACG AAC CAG GCA CAT 1020
341    V   I   T   M   P   P   P   I   H   A   K   Y   I   S   S   G   G   C   D   T   360
1021  GTT ATC ACA ATG CCA CCA CCA ATC ACT GCT AAA TAT ATA TCT AGT GGA GGT TGT GAT ACG 1080
361    K   E   N   N   S   V   A   K   N   I   S   M   P   S   G   M   F   S   Y   E   380
1081  AAG GAG AAC AAT TCT GTT GCG AAA AAC ATT TCA ATG CCA TCT GGA ATG TTC TCC TAC GAA 1140
381    E   L   S   K   A   T   G   G   F   S   E   E   N   L   L   G   E   G   G   F   400
1141  GAA CTT TCA AAA GCA ACT GGT GGA TTT TCA GAG GAG AAC CTT TTG GGA GAA GGC GGT TTC 1200
401    G   Y   V   H   K   G   V   L   K   N   G   T   E   V   A   V   K   Q   L   K   420
1201  GGA TAT GTT CAC AAA GGA GTG TTG AAA AAC GGG ACA GAA GTT GCG GTG AAG CAG CTG AAG 1260
421    I   G   S   Y   Q   G   E   R   E   F   Q   A   E   V   D   T   I   S   R   V   440
1261  ATT GGG AGC TAT CAA GGG GAA AGA GAA TTC CAA GCT GAG GTT GAC ACA ATC AGT AGG GTT 1320
441    H   H   K   H   L   V   S   L   V   G   Y   C   V   N   G   D   K   R   L   L   460
1321  CAT CAT AAG CAC CTC GTT TCA TTG GTT GGT TAT TGC GTT AAT GGA GAT AAA AGA CTC TTG 1380
461    V   Y   E   F   V   P   K   D   T   L   E   F   H   L   H   E   N   R   G   S   480
1381  GTT TAC GAG TTT GTT CCT AAA GAT ACC TTG GAG TTC CAC TTG CAT GAG AAC AGA GGA AGC 1440
481    V   L   E   W   E   M   R   L   R   I   A   V   G   A   A   K   G   L   A   Y   500
1441  GTG TTG GAA TGG GAA ATG AGG CTC AGG ATT GCT GTA GGA GCA GCA AAA GGA TTA GCT TAT 1500
501    L   H   E   D   C   S   P   T   I   I   H   R   D   I   K   A   A   N   I   L   520
1501  CTT CAT GAG GAT TGC AGT CCA ACT ATA ATT CAC CGT GAT ATC AAA GCA GCT AAT ATC CTT 1560
521    L   D   S   K   F   E   A   K   V   S   D   F   G   L   A   K   F   F   S   D   540
1561  CTA GAT TCC AAA TTT GAG GCA AAG GTC TCT GAC TTT GGA CTA GCC AAG TTT TTC TCA GAC 1620
541    T   N   S   S   F   T   H   I   S   T   R   V   V   G   T   F   G   Y   M   A   560
1621  ACC AAT TCA TCA TTC ACT CAT ATC TCT ACT CGA GTG GTA GGA ACT TTC GGA TAC ATG GCT 1680
561    P   E   Y   A   S   S   S   G   K   V   T   D   K   S   D   V   Y   S   F   G   V   580
1681  CCA GAA TAC GCG TCC AGT GGT AAA GTA ACT GAT AAA TCA GAT GTA TAT TCC TTT GGG GTC 1740
581    V   L   L   E   L   I   T   G   R   P   S   I   F   A   K   D   S   S   T   N   600
1741  GTG CTT CTA GAA CTC ATC ACT GGA CGT CCA TCA ATT TTC GCC AAA GAT TCT TCC ACA AAC 1800
601    Q   S   L   V   D   W   A   R   P   L   L   T   K   A   I   S   G   E   S   F   620
1801  CAG AGT TTA GTA GAC TGG GCG AGG CCA TTG CTT ACG AAA GCA ATC TCT GGA GAA AGT TTT 1860
621    D   F   L   V   D   S   R   L   E   K   N   Y   D   T   T   Q   M   A   N   M   640
1861  GAC TTT CTT GTA GAC TCA AGG TTG GAG AAG AAT TAC GAT ACA ACT CAG ATG GCA AAC ATG 1920
641    A   A   C   A   A   A   A   C   I   R   Q   S   A   W   L   R   P   R   M   S   Q   660
1921  GCT GCT TGT GCT GCT GCT TGC ATA CGC CAA TCA GCT TGG CTT CGG CCT AGA ATG AGC CAG 1980
661    V   V   R   A   L   E   G   E   V   A   L   R   K   V   E   E   T   G   N   S   680
1981  GTA GTA CGT GCT CTT GAA GGC GAG GTG GCC CTG AGA AAG GTC GAA GAG ACT GGG AAT AGC 2040
681    V   T   Y   S   S   S   E   N   P   N   D   I   T   P   R   Y   G   T   N   K   700
2041  GTG ACC TAT AGC TCT TCT GAA AAC CCG AAT GAC ATC ACA CCA CGG TAT GGA ACA AAT AAG 2100
701    R   R   F   D   T   G   S   S   D   G   Y   T   S   E   Y   G   V   N   P   S   720
2101  AGG AGA TTC GAC ACA GGT TCA AGC GAT GGT TAC ACT TCA GAA TAT GGA GTT AAC CCT TCT 2160
721    Q   S   S   S   E   H   Q   Q   V   N   T   *                                  732
2161  CAG TCG AGC AGT GAA CAT CAA CAG GTG AAT ACT TAG                                 2196
```
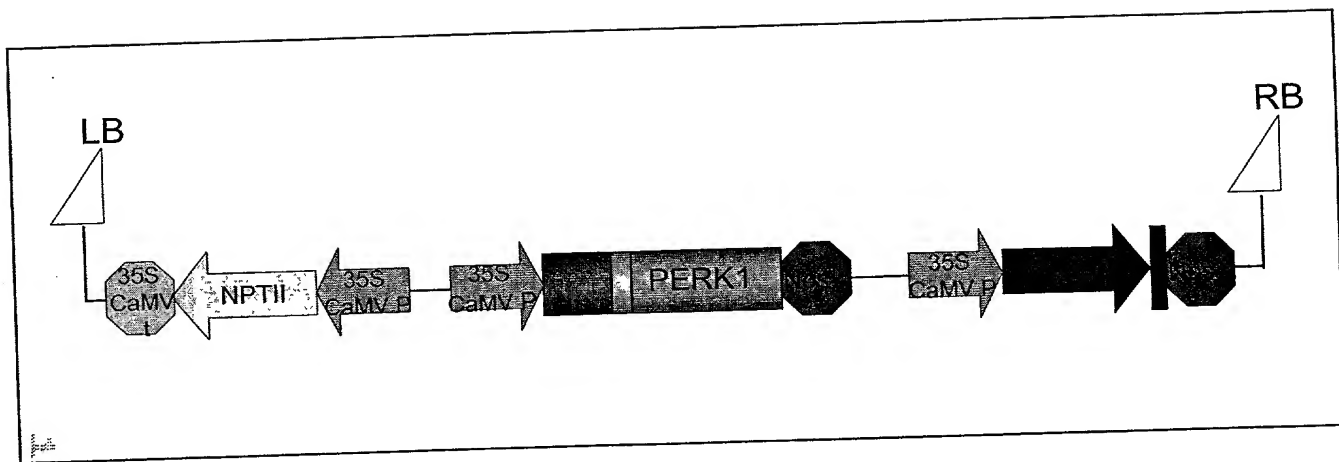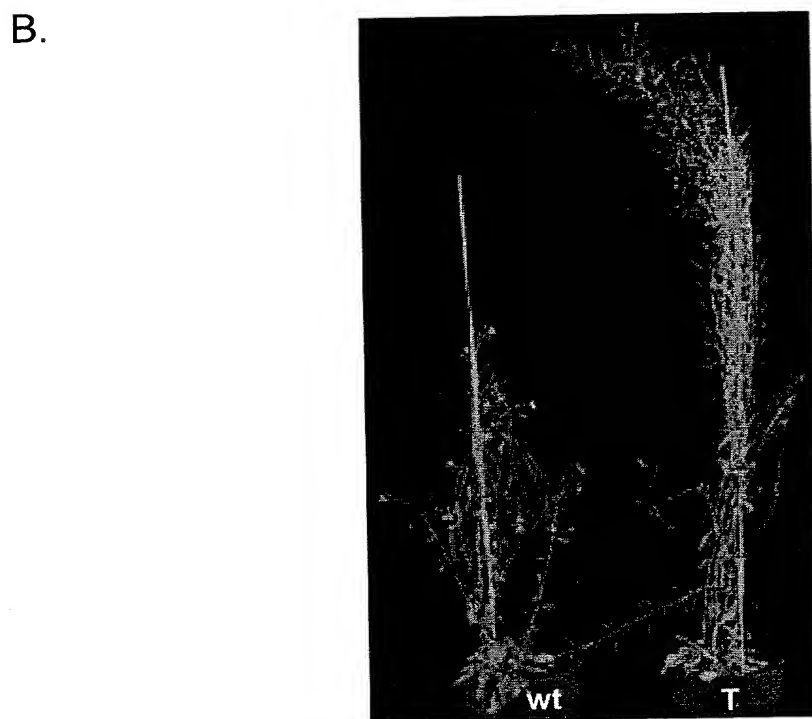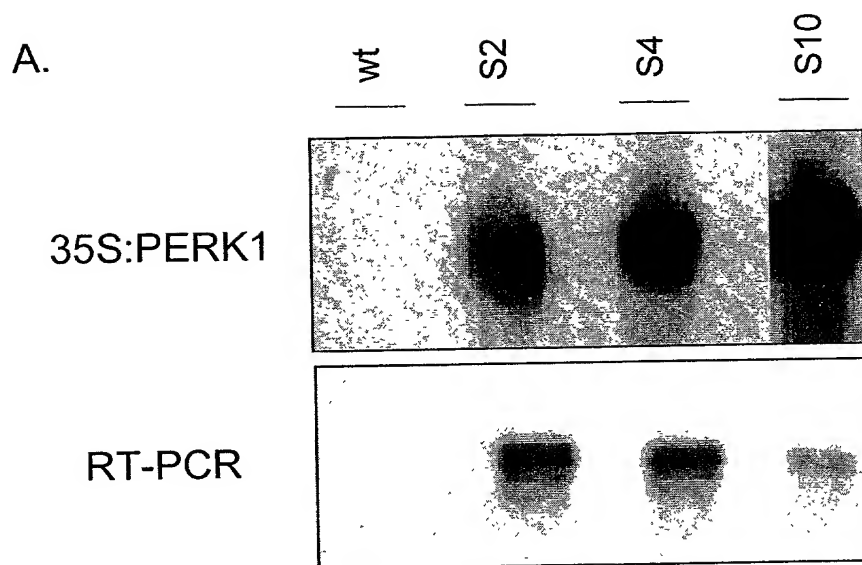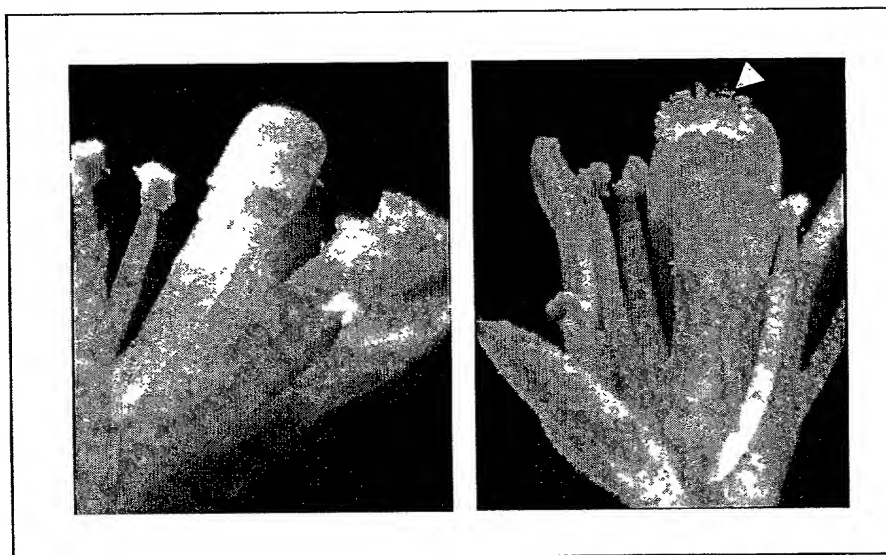
Figure 14B

Figure 15

A.

wt   S2   S4   S10

35S:PERK1

RT-PCR

B.

Figure 16

Figure 17

A.

	Wild-type	Overexpressing PERK1 Lines
Average seeds/silique	37	54
Average dry seed weight/plant	60.2 mg	106.2 mg

B.

Figure 18

Fig. 19